Chapter 3

XAFS EXPERIMENTS

This chapter will discuss the experimental procedures used to collect the XAFS data in this study. Most of the discussion here is applicable to general XAFS experiments. Special aspects for the XAFS study of dilute binary alloys, including the measurement of XAFS signals in fluorescence mode, the preparation of samples suitable for XAFS, and the control of high-temperatures during the XAFS measurement will also be discussed. A general description of the sample preparation techniques used in this study will also be given here, though further details for each of the systems studied will be given in later chapters.

3.1 Collection of XAFS Data

An XAFS measurement is a simple but careful measurement of the x-ray absorption attenuation coefficient $\mu_{\text{total}}$, as a function of x-ray energy for energies at and above an absorption edge. This is not particularly difficult for most samples studied by XAFS, especially since the advent of modern synchrotron sources with beam-lines dedicated to collection of XAFS data. There are a few subtleties, however, as will be discussed here. Figure 3.1 shows a schematic diagram of an absorption measurement.

3.1.1 The Incident X-rays

The first requirement for getting a good and reproducible measurement of $\mu_{\text{total}}$ is to have an intense x-ray beam. And since absorption measurements might be made on a wide range of central atoms and absorption edges, a broad energy range is desirable. These two requirement have come to mean using a synchrotron x-ray source, which gives both an intense x-ray beam and a broad energy range [63]. The x-ray beam coming from a synchrotron source is also well-collimated, so that divergence of the x-ray beam is small and the experiment can be placed well away from the source (so that other devices can be placed between the sample and source) without significant loss of beam intensity. A synchrotron also provides a polarized x-ray beam, which can
Figure 3.1: schematic lay-out of an x-ray absorption experiment, showing a double crystal monochromator, collimating slits, and positions of ion chambers relative to sample.

be exploited to select the XAFS from a subset of the local structure of the sample. The present work does not use the polarization of the x-ray beam.

To get useful and reliable fine structure details from an absorption measurement, the beam must have good energy resolution at the desired energy. This can be done (and has been done for measurements made in this work) using a double crystal monochromator. An x-ray beam of broad energy spectrum from the synchrotron source is incident on a monochromator, which allows only a narrow band of energies to be transmitted, while essentially all other energies are rejected. The monochromator is typically made from two nearly perfectly parallel pieces of single crystal silicon, cut to the same low index plane (\(\{111\}\), \(\{311\}\), and \(\{220\}\) are common).

X-rays are diffracted from the first crystal at an angle \(\theta\) (as shown in Fig.3.1) which depends on the x-ray energy, \(E_\gamma\), and spacing between crystal planes \(d\), according to the Bragg law,

\[
\theta = \sin^{-1}\left(\frac{hc}{2dE_\gamma}\right),
\]

where \(h\) is Planck’s constant and \(c\) is the speed of light. The second crystal is used to
re-diffract the now spatially and angularly separated x-ray beam, allowing only one set of discrete energies to emerge parallel to original incident beam. An x-ray beam with a very narrow spread of energies can be selected by using slits far away from the monochromator (and usually quite close to the sample). These slits will block x-rays which were diffracted by the second crystal at an angle slightly off-axis, and therefore with energies slightly different than the desired energy. Using this simple set-up, energy resolutions on the order of 1 to 3 eV, which are suitable for most XAFS experiments, can routinely be achieved.

A minor complication with a double-crystal monochromator is that the Bragg law allows harmonics (integer multiples of wavelengths) to be diffracted at nearly the same angles as the principle energy, and these harmonics must be eliminated before they enter the experimental chamber. This is achieved two ways. The second harmonics (which will be the brightest of the harmonics) can be fully eliminated by using monochromator crystals cut to a crystal plane that will not diffract the second harmonic. For silicon (with the diamond crystal structure), the rules for choosing planes which will forbid the second harmonic are that the indices $h, k$, and $l$ are either all odd, or are all even and sum to 4 [64]. The planes with $\langle hkl \rangle = \langle 111 \rangle, \langle 311 \rangle$, and $\langle 220 \rangle$ all satisfy these conditions. Third and higher harmonics are rejected by using the fact that each diffracted energy has some angular spread (due to divergence of the source beam and inherent disorder in the monochromator crystals). The higher energy beams of the harmonics generally have less angular spread than the fundamental. By slightly rotating the second crystal away from parallel with the first crystal (usually called “detuning” the monochromator, by moving away from the peak of its “rocking curve”) most of the intensity of the harmonics can be rejected, while the intensity of the fundamental remains substantially unchanged.

3.1.2 Measuring X-ray Intensities with Ion Chambers

The second important consideration in measuring an XAFS signal is to faithfully measure the intensity of a monochromatic x-ray beam. All x-ray intensities in this work were measured using ion chambers, a technology that is both simple and reliable. An ion chamber consists of two parallel plate electrodes with a uniform electric field and an inert gas between them. It measures the intensity of an x-ray beam by absorbing a fraction of the beam that passes through it, creating a photoelectron-ion pair. Inert gases (He, N$_2$, Ne, Ar, or Kr) are used because they will not recapture free electrons and because they can be mixed without affecting their properties. The
charged carriers collide with other gas atoms, creating daughter electron-ion pairs until the energy of the initially absorbed x-ray is distributed over many electron ion pairs. For inert gases, the ratio of incident photon energy and number of pairs created is fairly constant, with an effective ionization energy of 15 to 30 eV/pair. The charged carriers are swept away by an electric field (typically on the order of 100 V/cm) to the electrodes where they are collected and measured as a current. A field of this magnitude allows electron-ion pairs to be separated without recombining, and prevents the photo-electron from accelerating so much as it drifts toward the anode to cause avalanche collisions, which would make the chamber into a proportional counter, which has considerably worse signal-to-noise characteristics for fluxes typical in an XAFS experiment.

The current measured from an ion chamber is proportional to \( N_{\gamma} \), the number of x-rays absorbed per second and \( E_{\gamma} \), the energy of the x-ray, and is inversely proportional to \( E_{\text{gas}} \), the effective ionization energy of the gas, so that the current measured is (with \( e \) the charge on an electron)

\[
I_{ic} = \frac{2e N_{\gamma} E_{\gamma}}{E_{\text{gas}}}.
\]  

(3.2)

With typical values of \( E_{\text{gas}} = 30 \text{ V}, E_{\gamma} = 10 \text{ KeV} \), and a low but typical (especially for fluorescence measurements) flux \( N_{\gamma} = 10^6 \text{ photons/sec} \), the ion chamber current will \( I_{ic} \approx 100 \text{ pA} \). Such small currents need to be measured with electronic signal-to-noise levels better than \( 10^{-4} \), so that measurements of \( \mu_{\text{total}} \) will be limited by inherent sample characteristics and shot noise in the photon flux for weak signals (\( N_{\gamma} \lesssim 10^8 \text{ photons/sec} \)). Some subtle (but not formidable) electronic considerations must therefore be addressed to get a good signal from an ion chamber, including high-gain current amplification, acoustic pick-up, signal isolation, and transmission of signals. Commercially available current amplifiers (as from Keithley, for example) can approach these requirements for low-current signal-to-noise levels. We use such amplifiers for incident fluxes \( N_{\gamma} \gtrsim 10^8 \text{ photons/sec} \). For weaker signals (as is typical in the measurements of fluorescence signals discussed below), we use a low-noise current amplifier designed by Stern and built by the Stern research group over a decade or so of modifications [65, 66].

Ion chambers use the absorptive properties of inert gases to give their measured currents. For a chamber of thickness (or path-length of the x-ray beam) \( t \) and a total absorption coefficient of \( \mu_{\text{gas}} \), the number of x-rays absorbed per second, \( N_{\gamma} \) will be

\[
N_{\gamma} = N_I \left[ 1 - \exp \left( -\mu_{\text{gas}} t \right) \right]
\]  

(3.3)
where $N_I$ is the incident x-ray flux at the ion chamber. By mixing different inert gases together, the total value of $\mu^{\text{gas}}$ can be adjusted so that the ion chamber can absorb any fraction of the incident x-rays. An important consequence of using gases to sample the x-ray intensity is that the sampling fraction of ion chambers is energy-dependent. This arises from the inherent energy dependence of $\mu$ for each gas. Fortunately, as long as the ion chambers are not used at energies near the absorption edges of the gases inside, the absorption coefficients do not change dramatically over the 1 keV range of an XAFS experiment. More importantly, they are smooth and well-characterized functions of energies, so that tables of their values [67] can be used to theoretically correct for the energy-dependence of ion chamber response. Such corrections will be called “McMaster corrections” throughout this work, after the tabulated values used to estimate this effect.

### 3.1.3 Transmission Measurements of $\mu$

With the ability to produce a monochromatic x-ray beam with a small spread in energy, and the ability to measure the intensity of such a beam, the measurement of $\mu$ is nearly at hand. There are a few different ways to measure $\mu$. I will limit the discussion here to the transmission and fluorescence techniques, the two methods used for this work. The choice between which of these methods to use depends on the sample to be measured, and includes issues of signal-to-noise ratios, absorption by atoms other than the probe atom, and the ability to make uniform samples of the appropriate thickness.

A transmission experiment uses Eq. (2.1) directly, or nearly so. The incident x-ray beam intensity $I_0$ is sampled (as with Eq. (3.3)) to give a value $N_0$. Similarly, the x-ray beam transmitted through the sample $I_T$ is sampled to give a value $N_T$, and the total absorption coefficient is determined from the logarithm of their ratio

$$\mu_{\text{total}} x = \ln(N_0) - \ln(N_T),$$

where $x$ is the thickness of the sample. This is usually the simplest and least error-prone way to measure $\mu$, and works well for concentrated samples that can be made sufficiently thin and uniform. Still, there is the subtlety of optimizing the signal-to-noise level in a transmission experiment, which depends on the absorption fraction of each of the two ion chambers, and on the absorption fraction of the sample. These are not difficult to determine and have been described by many writers [31, 68], so that the results will simply be used here.
Typically 15% of the incident x-ray intensity is absorbed in the $I_0$ chamber, and a mixture of inert gases are chosen to fill the chamber and have $\mu_0^{\text{gas}} t \approx 0.15$. The $I_T$ chamber is similarly filled with a mixture of inert gases so that $\mu_0^{\text{gas}} t \approx 2.5$, allowing 90% of the x-rays that enter it to be absorbed. It is important to prevent $\mu_0^{\text{gas}} t$ from being too large, as this would cause most of the x-rays to be absorbed in a small region of the ion chamber, where the electric fields of the ion chamber plates can be distorted. With $\mu x \approx 2.5$, roughly 70% of the x-rays are absorbed in the first half of the ion chamber, so that the entire chamber is used to absorb all the x-rays.

The optimum thickness of the sample is generally taken to be $\mu x \approx 2.5$. This value takes into account only theoretical signal-to-noise levels, and ignores practicalities like being able to make a uniform sample of this thickness. A non-uniform sample can cause rather severe problems in the quality of the measured $\mu$, most of which are known as “thickness effects”. The simplest of these comes from having pinholes in the sample, so that part of the transmitted beam is not attenuated at all by the sample. This gives an unwanted addition to the background absorption signal, and deteriorates the resulting XAFS spectra.

### 3.1.4 Fluorescence Measurements of $\mu$

The results of the signal-to-noise considerations in the previous section were based only on the total $\mu$ of the sample, and gave no consideration to how this total was achieved by the sample. For many classes of samples, the total absorption is dominated by atoms other than the probe atom in the study. This can happen either because the probe atom is a minority component of the sample or because it is a much lighter atom than the rest of the atoms in the sample. In either of these cases, the change in $\mu_{\text{total}} x$ of the sample at the edge step will be much less than 1. This unwanted absorption will degrade the effective signal-to-noise ratios for the edge-step (which is the important of an XAFS measurement).

For sufficiently dilute samples, with the absorption from the edge of interest significantly less than the total absorption ($\mu_0 \ll \mu_{\text{total}}$), the transmission technique of measuring absorption can be abandoned in favor of the fluorescence technique. In this method the absorption is measured by sampling the x-rays fluoressed in the de-excitation of the core-hole. The fluorescent energies have discrete values, typically 10% lower in energy than the absorption edge energy. The probability of de-excitation by fluorescence (the major competing de-excitation channel is the Auger process) is called the fluorescence efficiency, $\epsilon_f$, and increases with increasing fluorescence energy.
Fluorescent x-rays are emitted from the absorbing atom isotropically, and so are sent from the sample in essentially all directions. This puts a premium on solid angle subtended by the fluorescence detector, so as to capture as many of the fluoresced x-rays as possible. To increase the number of fluoresced x-rays making it into the fluorescent ion chamber, the sample is rotated in the x-ray beam. It is usually very nearly optimal for the sample to make a \(45^\text{o}\) angle to both the incident x-ray beam and the plane of the fluorescence ion chamber (as shown in Fig. 3.1). For this geometry, and with a sufficiently thick sample (so that the transmission is negligible), the number of x-rays entering the fluorescent x-rays entering the fluorescence ion chamber is

\[
N_f(E_f) = \epsilon_f F_\Omega N_0(E) \left\{ \frac{\mu_\alpha(E)}{\mu_{\text{total}}(E) + \mu_{\text{total}}(E_f)} \right\},
\]

where \(F_\Omega\) is the fraction of the \(4\pi\) solid angle filled by the detector and \(N_0\) is the number of x-rays incident on the sample. As above, \(\mu_\alpha\) is the absorption coefficient for the edge of interest and \(\mu_{\text{total}}\) is the absorption coefficient for the entire sample. The dependence on \(\mu_{\text{total}}\) at both the changing incident energy \(E\) and the constant fluorescent energy \(E_f\) reflects the dependence on getting an x-ray of energy \(E\) into the sample and an x-ray of energy \(E_f\) out of the sample.

If the absorption of the sample is dominated by the edge of interest, then \(\mu_\alpha \approx \mu_{\text{total}}\) and the term in braces in Eq. (3.5) is essentially constant, so that the fluorescence signal does not vary with energy. But if \(\mu_\alpha \ll \mu_{\text{total}}\) then it is the denominator of the term in braces in Eq. (3.5) that is essentially constant, and \(N_f \sim \mu_\alpha(E)\). Thus for sufficiently dilute samples, the fluorescent signal is directly proportional to the absorption coefficient of the edge of interest.

The question then arises of how to decide whether it more favorable to measure the absorption coefficient by transmission or fluorescence for a given sample. Again, I defer such rather straightforward signal-to-noise calculations to the standard references [30, 31], and just state the results here. In the absence of sources of noise other than photon counting statistics, fluorescence and transmission measurements of \(\mu\) will be equally favorable when

\[
\mu_\alpha \approx \epsilon_f F_\Omega \mu_{\text{total}},
\]

where, as before, \(\mu_\alpha\) gives the absorption due to the edge of interest, \(\mu_{\text{total}}\) gives the total absorption. Since \(F_\Omega\) is typically 10\% of the \(4\pi\) solid angle, and since \(\epsilon_f\) ranges from 0.1 to 1.0, it is usually the case that fluorescence is favorable for concentrations of the probe atom less than several per cent.
When making a fluorescence measurement there is an additional complication in that the sample can scatter x-rays into the fluorescence chamber. Both elastic or inelastic (primarily Compton) scattering occur, but the change of energy in Compton scattering is rarely as much as the difference between edge excitation and fluorescence energies. The elastic and nearly-elastic Compton scattering can usually be prevented by using x-ray filters and Soller slits [30]. An x-ray filter is simply a piece of material that has an x-ray absorption edge between the absorption edge and the fluorescent energy of the atom of interest. This will allow the fluorescent x-rays to be passed, while the elastically and nearly-elastically scattered x-rays are passed. For most K-edges, a suitable filter is made of atoms with $Z - 1$ from the element of interest. Soller slits, which make a kind of filter commonly used to eliminate diffuse scattering by preferentially passing x-rays from a certain place, are used to prevent the fluorescence of the atoms in the x-ray filter (which absorb x-rays, and so fluoresce too) from entering the fluorescence ion chamber. The use of x-ray filters and Soller slits can dramatically increase the effective counting rate of the fluorescence ion chamber by eliminating a large portion of the x-rays that do not come from the desired fluorescence of the probe atom.

3.1.5 Practical Strategies for the Collection of XAFS Data

After the method used to measure $\mu_{\text{total}}$ has been chosen, and the mixture of gases in the ion chambers has been selected, there are still some considerations for getting the best XAFS measurement. These are predominantly strategies for getting sufficiently good data in as little beam-time as possible. I will outline such considerations here.

The first important consideration is to make certain that the x-ray beam is sufficiently monochromatic. Detuning the monochromator to eliminate third and higher harmonics is essential for lower energy edges, but becomes less important as the absorption edge increases. The Ag K-edge, for example, has a third harmonic at $\sim 76$ KeV, where the intensity of the NSLS synchrotron is very low. It is therefore important not to detune higher energies too much. The energy resolution should also be optimized to give a monochromatic beam. This can be done by changing the vertical slit size of the beam just in front of the $I_0$ chamber until the resolution of the measured absorption edge (which is affected by not only the resolution of the incident radiation but also the intrinsic resolution of the excited core-hole) cannot be improved. It is not necessary to close the slit size any further than this, as the resolution will not be improved and the incident flux will be reduced.
The second consideration is to make sure the cross-section of the x-ray beam is completely and only on the sample, but that it is hitting as much of the sample as possible. Careful alignment is crucial for proper normalization of the transmitted (or fluoresced) intensity with the incident intensity. If the \( I_0 \) chamber sees a different part of the x-ray beam than the sample or the \( I_T \) (or \( I_f \)) chamber sees, then the reason for using the \( I_0 \) chamber (which is to sample the actual beam hitting the sample) is lost. The beam may have spatial variations in its intensity (this is especially true for “glitches” and harmonics) and these must completely canceled. Usually horizontal slits can be adjusted to optimize the photon flux on the sample.

The third consideration is to choose the scanning range, step sizes, integration times, and the number of scans taken. It is always a good idea to take more than one scan, and, if possible more than two. This will help track down any systematic distortions in the measurement. The total integration time is also important as this will give the total flux counted, which determines the shot noise of the measurement. For transmission measurements, a total count of \( \sim 10^8 \) photons/point is desired as the random noise will by \( \sim 10^{-4} \). For fluorescence measurements, such luxurious counting rates are rare, and an effective counting rate (described below) of \( \sim 10^6 \) photons/point is acceptable. The step sizes of the individual scans should also be chosen carefully. In the pre-edge region it is important to get a few (\( \sim 10 \)) points, separated by 20 eV or so. In the edge-step region, the step size should be reduced to 1 or 2 eV. Above the edge, it is best to space the points evenly in \( k \), as this is how the data will later be analyzed. In view of the information content of XAFS (see section 4.1) a spacing of \( 0.05 \AA^{-1} \) is a good choice. In principle this will allow \( R \)-space features to be seen out to \( \sim 30 \AA \), and in practice will allow interpolations of the data without affecting data below 10 Å.

The fourth consideration is to make sure the collected data is free from systematic errors, especially “glitches” in the data. These are sharp changes in the measured \( \mu_{\text{total}} \) that are very localized in energy, typically a few eV wide. They are caused by non-linear changes in the incident beam intensity profile that do not fully cancel when compared to the intensity of the beam transmitted through (or fluoresced from) the sample. The presence of glitches usually means that the first two considerations above were not properly addressed. By adjusting the beam alignment these glitches can usually be sufficiently reduced. If glitches cannot be eliminated in this way, it is likely that the sample is inhomogeneous (or crystalline, so the glitches are Bragg peaks). If this is the case, then either making a more homogeneous sample or trying other regions of the sample should minimize the glitches.
If fluorescence measurements are done, then an additional consideration is the choice of x-ray filter type (this is usually trivial) and filter absorption thickness. This latter choice is best made by testing different filter thicknesses and picking the one which maximizes the effective signal-to-noise ratio. Since the signal goes as the size of the edge-jump $\Delta \mu x$ and the noise goes as the square-root of the total $\mu x$, the effective signal-to-noise ratio which should be optimized is

$$\left( \frac{S}{N} \right) = \sqrt{\frac{N_f(E_0 + 100) - N_f(E_0 - 100)}{N_f(E_0 + 100)}}.$$  \hfill (3.7)

Here $N_f(E_0 + 100)$ and $N_f(E_0 - 100)$ are the fluorescence signals just above and just below the edge jump, respectively. While optimizing the filter thickness, it is also a good idea to try different spatial positioning of the ion chamber (so that the Soller slit may be optimally aligned) and to try different rotation angles of the sample.

### 3.1.6 Conversion of Raw Data to $\mu(E)$

After the absorption data have been collected it is converted to experimental absorption data. Here the quality and reproducibility of the absorption data are assessed, and the extent of systematic errors in the absorption data is determined. The most important systematic error comes from the sharp glitches discussed in the previous subsection. If such glitches cannot be eliminated while the data is being collected, and if they are well-localized in energy, then they can easily be removed from the data. Many suggestions for the best method of removing glitches have been made (including replacing the absorption datum with the polynomial interpolation from the surrounding data points to the energy value) but in view of the information content of the XAFS signal (discussed in chapter 4), all methods are essentially equivalent to simply removing the offending data point.

Other systematic errors in the measured absorption can be more difficult to detect after the data are taken, but can significantly affect the measured $\mu$, especially its amplitude. It is outside the scope of this work to discuss all of these in detail, as they are well-covered by the standard XAFS references \cite{30, 32, 31}, but I will mention two important systematic errors in the measurement of $\mu$.

As mentioned above, ion chambers (like everything else) have an energy-dependent absorption coefficient. This will cause an energy-dependence to the measured $\mu$ for the sample. For transmission experiments the energy-dependence of the $I_0$ and $I_T$ chambers cancel nearly completely. Since the logarithm of the ratio of these two signals is taken, most of the difference in energy-response of the two chambers gives a
constant offset to the measured $\mu$. The remaining difference gives a small and smooth energy-dependence to $\mu$ which is easily removed in the background removal.

For fluorescence measurements the energy response of the two chambers do not tend to cancel, because the $I_f$ chamber measures the intensity of a fixed energy (the fluorescence energy is a characteristic of the atom and does not depend on the excitation energy), while the $I_0$ chamber measures the intensity of a beam whose energy varies. In addition, the logarithm of the intensities is not taken for fluorescence measurements, so that even the exponential decay of the $I_0$ response is included in $\mu$. Fortunately, the energy-dependent absorption properties of the inert gases is well-known [67], and can be easily corrected for a given energy range and $I_0$ gas. It turns out that the nearly exponential decay of atomic absorption coefficients in the hard x-ray region gives a correction to the $\sigma^2$ term in the XAFS equation of Eq. (2.53). The value of this $\sigma^2$ correction is easily calculated by the ATOMS program [69].

Another correction that can be important for samples measured in fluorescence is the absorption feedback inherent in Eq. (3.5) for samples which do not have $\mu_\alpha \ll \mu_{\text{total}}$. The result of this condition is to measure a smaller $\mu$ than the true value, which reduces the XAFS amplitude. To first order, the reduction is by a constant amount, although there can be important energy-dependent contributions too. As for the ion chamber effects above, these systematic reductions in amplitude (in the form of both an overall reduction factor, and an additional $\sigma^2$) are easily calculated by the ATOMS program for a given given probe atom and sample composition.

3.2 XAFS Measurements at Low and High Temperature

In order to make the temperature-dependent XAFS measurements for this work, simple and reliable controls of the sample temperature during the XAFS experiment are essential. Low-temperature (below room temperature) measurements were made using a commercially available Displex refrigerator from Air Products, which uses a helium compressor to cool a metal cold-finger (onto which the sample can be mounted) down to $\sim 10$ K. Higher temperatures (up to room temperature) are stabilized by a heating element at the cold-finger which is turned on and off to control the temperature by using a feedback process. Either a thermocouple wire or diode can be used to convert the temperature at the cold-finger (or better still, at the sample itself) into a voltage signal for the controller to maintain.

The metal foil samples measured in this work were mounted by pressing them between a Cu block and a Cu plate, both of which had holes (roughly 5 mm $\times$ 15 mm
in cross-section) for the x-rays to pass through. Five samples could be mounted simultaneously, greatly increasing the efficiency of the measurements. The Cu block was attached to the refrigerator using a threaded Cu bolt, which gave sufficient thermal contact between cold-finger and sample. Thermal isolation of the sample (so as to reduce the thermal load on the refrigerator) was provided by a combination of a radiation shield of Aluminized mylar, and a mechanical vacuum (typically 10 to 100 mTorr) in which the sample, cold-finger, and radiation shield were kept. The metal vacuum shroud which housed the sample and cold-finger had large holes cut in it, with kapton windows epoxied over the holes. The kapton windows maintained the mechanical vacuum and let x-rays pass.

In order to make high-temperature XAFS measurements, a furnace which could get to and maintain sample temperatures up to 1000 K was needed for this work. The design of the furnaces used in this work was based on earlier work done by Stern, and modified for this work to increase the temperature range. In the end, two different furnaces were used. One could reach 1000 K, and could hold three sample, but could only be used for fluorescence measurements. The other could reach 900 K, hold five samples, and be used for either transmission or fluorescence measurements.

The furnace is a copper block mounted on stainless steel tubes which can be placed in the same vacuum shroud used for the low-temperature measurements. Heaters components from soldering irons (Ungar part #UTC-233) were held against the Cu block, and wired in parallel. Each heater was nominally 100 Ω at room temperature. The fluorescence-only heater had six such heaters. The transmission furnace used only four heaters (the space where the other two heaters had been being empty so that x-rays could be transmitted through holes in the Cu block of the furnace). As for the low-temperature measurements, the samples were held in place by pressing them between the Cu block and a Cu plate with holes cut in it for x-rays to pass through. In order to reduce the thermal load on the furnace due to radiation of oxidized Cu plates, they were electro-plated with Ag. The thermal load on the furnace was further reduced by a mechanical vacuum and a radiation shield of aluminum foil (~ 9 μm thick, which absorbs very few x-rays at energies above 10 KeV).

The temperature on the samples was measured using type K thermocouple wire, with the two wires spot-welded together. The temperature was monitored by placing this wire between Cu block and plate with the samples, not more than 20 mm from the sample. Usually two such thermocouple wires were held at slightly different spots on the furnace. The temperature measured by the two wires seldom differed by more than 4 K. The temperature was controlled with a temperature controller
from Omega Corp., which turned on and off the power to the heaters (when on, 10 A was put through the heaters) using a rate-controlled feedback on the thermocouple voltage. With proper tuning of the controller’s feedback parameters, and starting from a temperature \( \sim 100 \text{ K} \) below the target temperature, any temperature could be routinely be achieved and held stable in a matter of minutes.

### 3.3 Sample Preparation

All samples measured in this study were made using essentially the same techniques. The few variations will be explicitly discussed. As discussed in section 3.1, the only important constraints on a sample for use in an XAFS experiment is its thickness. For the measurement of the minority component in a binary alloy, the absorption cross-section is best measured in fluorescence. In this case, the thickness of the sample, \( x \), is not critical, as long as \( x \gtrsim \mu^{-1} \). For most samples in this study, \( \mu^{-1} \gtrsim 10 \mu\text{m} \), and there was no difficulty in making a single foil out of the metallic sample of suitable size (typically \( 1 \text{ cm} \times 10 \text{ cm} \times 100 \mu\text{m} \)).

Metallic elements (in the form of millimeter size shot) were weighed to the proper mixtures (typically 2 and 5 at.\% ), and mixed. The mass of the host material was typically on the order of 10 gr, while that of the impurity was on the order of 0.1 gr. Host elements (Ag, Au, and Pb) were nominally 99.99999\% pure. The purity of the elements used as the impurity elements was never less than 99.9\%. The elements were placed in a quartz tube (it was found to be easier and more foolproof to weigh the host element inside the tube, and to add the impurity to the host until the desired total mass was obtained.), and hermetically sealed under mechanical vacuum (the pressure inside the sealed tube was on the order on 10 mTorr).

If the melting temperature of either of the elements was low enough (~ 700 K), the sealed container was heated over a gas flame until the element(s) melted. The melt was shaken until homogeneous. For the common situation in which the impurity would melt but not the host would not, (Ag and Au could not be melted over a Bunsen burner), the sample was shaken over the flame until the impurity completely wetted the solid host.

The sealed tubes were then placed in a furnace, and heated to at least 50 K above the melting temperature of the pure host. The temperature inside the furnace was monitored and controlled using a type K thermocouple wire wrapped around the quartz tube holding the sample. The samples were kept melted for times ranging between 24 and 48 hours. After that time the samples were quenched from the melt...
by dropping the quartz vials into water. Temperatures cool enough to handle were always achieved in a matter of 10 to 20 sec. As long as the quartz tube did not crack in the quenching process (which meant the samples were useless, and which happened only twice), the tubes were then broken, and the ingot of alloy was mechanically hammered and rolled to the desired thickness. For several alloys (especially those in Ag) became work-hardened while still several hundred μm thick. These foils were then resealed in quartz tubes and annealed (typically at 800 K for 24 h). After annealing, the Ag-based alloys could be further rolled to a suitable thickness.

The mechanically-pumped pressure (∼ 30 mTorr) of the vacuum shroud was not sufficient to ensure that the samples would not oxidize at high temperatures. To prevent such oxidation, the samples were coated with aluminum after being mechanical rolled to the desired thickness. The coating was done by evaporating a small amount of aluminum in a vacuum of 10 to 100 μTorr in which the foils were placed. The foil was flipped over, and the process repeated so that both sides of the foils were coated. Since the diffusion rate of oxygen through aluminum is quite low, this coating essentially prevented any oxygen from getting to the lead, silver, or gold hosts. But, since aluminum is miscible with both silver and gold, this treatment was insufficient, and the aluminum coating was itself intentionally oxidized by placing the aluminum-coated foil over steam. The real danger was not so much oxidizing the host but the impurity. For the lead-based alloys, an oxidized surface layer of lead would not have been too serious, since the lead oxide layer would have remained at the surface and would have prevented oxygen from passing through it. But silver and gold do not themselves react with oxygen (or aluminum, or much of anything for that matter), and are in fact fairly transparent to oxygen. The danger of oxygen getting to the impurity was therefore much greater in these alloys, and the aluminum oxide coating was important in keeping oxygen away from the impurity atoms. Mössbauer measurements on Sn impurities in each of the three hosts [24, 25] show no evidence of SnO₂ in well-coated samples, even though the Mössbauer effect is quite sensitive to low concentrations of SnO₂ that is easily recognizable by a signature line-shift. In addition to this evidence against oxidized samples, the XAFS measurements themselves on the various systems, show no evidence for oxygen near-neighbors to the impurity atoms.
Chapter 4

ANALYSIS OF X-RAY ABSORPTION FINE STRUCTURE DATA

This chapter will discuss the analytic techniques used in this study to extract information about the local structure around the absorbing atom from an experimental XAFS spectrum. This is not meant as a complete discussion of XAFS analysis. The reader is referred to any of the several review articles [30, 31, 70] for more detailed discussion and further references on the standard techniques of XAFS analysis. Some aspects of the analysis techniques used in this study are sufficiently novel, however, that a fairly detailed discussion of them will be given. Many of the details of these relatively new techniques have been given elsewhere [71, 72]. The discussion here will be more complete.

An x-ray absorption spectrum contains both a smoothly varying background and fine-structure oscillations which depend on a combination of x-ray scattering terms and local structure terms. Because of such different and complicated contributions to the measured spectrum, the analysis of XAFS data is not trivial and the use of computers (and programmed algorithms) is essential. Aside from utility programs to manipulate and display data, three computer programs were used extensively to analyze the XAFS data in this study:

FEFF which calculates a theoretical XAFS spectra given a spatial configuration of atoms, based on the curved-wave, multiple-scattering expansion of the XAFS equation due to Rehr, Albers, and co-workers.

AUTOBK which separates the experimentally measured absorption into background and XAFS $\chi(k)$.

FEFFFIT which modifies the theoretical spectra from FEFF until a best fit is found to the XAFS data.

FEFF was written by John Rehr and his students Jose Mustre de Leon, Steven Zabinisky, and Alexei Ankudinov in collaboration with R. Albers [54, 73, 74, 75, 76], and
has been shown to be a flexible and fast way to generate reliable predictions of XAFS spectra on a wide variety of experimental systems [77, 78, 79]. Here I will discuss aspects of the application of FEFF to XAFS analysis.

AUTOBK and FEFFIT were written by me, with many important suggestions coming from Peter Liščič, Bruce Ravel, John Rehr, Edward Stern, Yizhak Yacoby, and Steven Zabinsky. These programs have proved to be useful and reliable for XAFS analysis on many systems [79, 80, 76]. The procedures, algorithms, and approximations of these programs will be discussed in detail in this chapter. AUTOBK and FEFFIT have been incorporated, along with the very useful program ATOMS (which generates an input file to FEFF with atomic coordinates, and calculates several XAFS amplitude corrections given a space-group description of a crystal) written by Bruce Ravel and a few utility programs, into the UWXAFS3.0 analysis package1.

4.1 The Information Content of XAFS Data

Before beginning a discussion of XAFS analysis, I will briefly discuss the information content in an XAFS measurement. This is important in getting reasonable and reliable answers from any analysis. Topics from signal analysis and information theory, originally developed for studying how information could best be sent along a transmission line, will be applied here to the topic of how information about a solid can be transmitted through XAFS data.

The information content of a signal is partly a matter of perception, which is to say that you have to know what you’re looking for in a signal. One can imagine observing the waveform on a telephone receiver and being unable to get any information from the signal using even the most complex analysis of the waveform, while a child who understands that it is an audio signal containing human speech in a particular language can easily receive large amounts of information. In an XAFS experiment, we are looking for information about the local structure around the central atom that is contained in the fine structure oscillations of the absorption coefficient. The XAFS signal is easily related to the information we want by a Fourier transform, so that basic concepts of information theory [81] can be readily applied to the interpretation of XAFS data.

---

1 As part of this package, these programs have been copyrighted by the University of Washington and a license is required for their use. The interested reader should contact Edward Stern for more information about the availability of the programs that make up the UWXAFS3.0 package and the applicability of these computer programs for general XAFS analysis.
A fundamental result of information theory (given by Shannon, as described by Brillouin [81]) is the sampling theorem which is essentially a restatement of the standard uncertainty relation for conjugate Fourier components. For a signal with time duration $\Delta t$ this theorem gives that two frequencies cannot be distinguished if they differ by less than $\delta \omega = \pi/\Delta t$. Since only a finite frequency range (or bandwidth) can be sent and sampled for a real signal, there is a finite limit on the number of distinguishable frequencies measurable in a signal. And since the information in a signal is transmitted through the time-duration of its frequencies, there is an upper limit on the amount of information in a signal. This upper limit on information can be quantified as the maximum number of independent points in the signal, given by the simple relation

$$N \approx \frac{\Delta t \Delta \omega}{\pi},$$

(4.1)

where $\Delta t$ is the time duration of the signal and $\Delta \omega$ is the measurable frequency range.

For signals with small amounts of information, the approximate nature of Eq. (4.1) would be best eliminated. This can be done fairly simply by looking at the Fourier series upon which the sampling theory is based. The interpretation from this approach is that the information is taken as the values of the Fourier coefficients, which come in pairs, spaced at intervals in $\omega$ of $\pi/2 \Delta t$. This gives the amount of information a clearer, more useful definition and immediately leads to

$$N = \frac{\Delta t \Delta \omega}{\pi} + 1,$$

(4.2)

where the 1 represents the constant term in the usual Fourier series expansion term, and corresponds to there being 1 coefficient at $\omega = 0$ and pairs of coefficients at all non-zero multiples of $\pi/2 \Delta t$. For the interpretation of XAFS, the standard XAFS equation, Eq. (2.43), is interpreted to be similar to a Fourier transform with conjugate variables are $k$ and $2R$ so that Eq. (4.2) becomes

$$N = \frac{2 \Delta R \Delta k}{\pi} + 1,$$

(4.3)

where $\Delta k$ is the measured (and useful) $k$-range for the data and $\Delta R$ is the $R$-range for which the information about the local structure is desired. For the analysis of signals that excludes the origin (i.e., for high-frequency or high-$R$ analysis), the “+1” term is inappropriate, and the information consists purely of pairs of Fourier coefficients. As pointed out by Stern [82], and based on arguments of correlation between successive points in the Fourier transform of a signal, the number of independent points can be
written as

\[ N = \frac{2\Delta R \Delta k}{\pi} + 2, \tag{4.4} \]

if the \( R \)-range is chosen so both endpoints are chosen at non-zero multiples of \( \pi/\Delta k \). Simple arguments for this expression are that the Fourier components come in pairs, and that an infinitesimally small \( R \)-range must still contain the information of one independent point. The “+1” formulation of the information content of an XAFS signal will be used in background removal (where the information at \( R = 0 \) is used), and the “+2” formulation (with carefully chosen endpoints, so that they are as close as possible to integer multiples of \( \pi/\Delta k \)) will be used in the structural analysis of the XAFS signal.

A corollary to the sampling theorem which is easily seen from the interpretation of information as Fourier coefficients is that two independently measurable frequencies can never be closer than \( \delta \omega = \pi/\Delta t \). In order to get the maximum amount of information from a signal, then, the measured frequencies must be evenly distributed. The consequence of this for the interpretation of XAFS signals is that information is evenly spaced in \( R \), without regard to where the truly interesting physics may be. This even distribution of information is generally considered a detriment to XAFS analysis, as it puts a strict limit on the number of measurable parameters at high-\( R \), where there are many more possible structural parameters that can influence the data. But, as will be seen in section 4.3.2, this even spacing of information does have a benefit for separating the XAFS \( \chi(k) \) from the atomic background absorption.

### 4.2 The XAFS Fourier Transform

Like most XAFS analysis procedures, the methods described here are based on the Fourier transform. The form of the Fourier transform is somewhat specialized to XAFS analysis, and the modifications made for XAFS analysis will be explained here. It might be argued that the form used is arbitrary and is not based on any physical or analytic principles. I won’t dispute such allegations of capriciousness, and would agree that seeking for a better representation could be worthwhile. But the historical form of the XAFS Fourier transform [70] does have some advantages (such as not depending at all on any knowledge of the system under study and the ability to emphasize different parts of a spectrum), and will be used exclusively in this work.

As mentioned in the previous section, and as suggested by the basic XAFS equation in Eq. (2.43), XAFS data is usually analyzed by using a Fourier transform with conjugate variables \( k \) and \( 2R \). Since there is only a finite \( k \)-range over which reliable
data can be measured, some sort of filtering of the data must be done — the Fourier transform of real data cannot extend to \( k = \infty \). This is usually accomplished with a window function \( W(k) \), discussed below. Finally, because the amplitude of XAFS \( \chi(k) \) data tends to decay with \( k \) (due to decay of back-scattering amplitude and disorder in the sampled structure), it is customary to use an additional weighting term of the form \( k^w \), with \( w = 1, 2, \) or 3. With all these pieces in place the XAFS version of the Fourier transform is then

\[
\tilde{\chi}(R) = \text{FT}[\chi(k)] = \frac{1}{\sqrt{2\pi}} \int_0^\infty k^w \chi(k) W(k) e^{i2kR} dk. \tag{4.5}
\]

The XAFS in \( R \)-space is written as \( \tilde{\chi}(R) = \text{FT}[\chi(k)] \) even though the conjugate to \( k \) is \( 2R \) not \( R \) — the Fourier transform must be mapped from \( 2R \) to \( R \).

The form of Eq. (4.5) is essentially useless for the manipulation of real data. A discrete version is needed, which will readily allow the Fast Fourier Transform algorithm [83] to be used. A discrete \( k \)-space grid of \( \delta k = 0.05 \text{ Å}^{-1} \) is used for all data analyzed in this work. Typical array sizes used are \( N = 2^9 = 512, N = 2^{10} = 1024, \) or \( N = 2^{11} = 2048, \) which are chosen because the Fast Fourier transform is fastest when the array size is a power of 2, and data \( k \)-ranges often exceed \( N = 256 \) points. Arrays are filled with zeros for \( k \)-values past the range of measured data. The \( R \)-space grid is \( \delta R = \pi/N \delta k \), and we write \( k_n = n \delta k \) and \( R_m = m \delta R \). The discrete form of the XAFS Fourier transform is

\[
\tilde{\chi}(R_m) = \text{FT}[\chi(k_n)] = \frac{\delta k}{\sqrt{\pi N}} \sum_{n=1}^N \chi(k_n) W(k_n) k_n^w e^{2\pi i m n / N}. \tag{4.6}
\]

All Fourier transforms discussed in this work will be of this form, and \( \tilde{\chi}(R) \) will be taken to mean a discrete Fourier transform of weighted and windowed \( \chi(k) \) data.

The Fourier transform window function, \( W(k) \), is used to smooth out the data while maintaining some peak separation of the \( R \)-space data. There are many functional forms of the window [84], which vary in their abilities to smooth data, reduce ripple, and maintain peak separation. The relative merits of the different window functions are hard to measure quantitatively, and end up being a matter of taste. The analysis programs of the UWXAFS3.0 package described in this chapter allow (at last count) 8 different forms of windows. But to be consistent, all data shown here will be done using the Hanning window function, a widely used form with a long history in XAFS analysis. This window has parameters \( k_{\text{min}}, k_{\text{max}}, \delta k, \) and a
functional form

$$\begin{align*}
W(k) &= \begin{cases} 
\sin^2 \left( \frac{\pi (k - k_{\text{min}} + dk/2)}{2dk} \right), & k_{\text{min}} - dk/2 \leq k < k_{\text{min}} + dk/2 \\
1.0, & k_{\text{min}} + dk/2 \leq k \leq k_{\text{max}} - dk/2 \\
\cos^2 \left( \frac{\pi (k - k_{\text{max}} + dk/2)}{2dk} \right), & k_{\text{max}} - dk/2 < k \leq k_{\text{max}} + dk/2 
\end{cases}.
\end{align*}$$

(4.7)

All data analyzed in this work will be done with this window, with typical values for $dk$ of \(\sim 1.5 \text{ \text{Å}^{-1}}\).

4.3 The Isolation of $\chi(k)$

The first step in the analysis of XAFS data is to convert the measured absorption data $\mu(E)$ (after appropriate evaluation of the quality of the data, and, if needed, the application of the corrections discussed in section 3.1.6) into the XAFS $\chi(k)$, which is usually defined (also see Eq. (4.9) below) as

$$\chi(E) = \frac{\mu(E) - \mu_0(E)}{\mu_0(E)}.$$  

(4.8)

$\chi(k)$ is formed from $\chi(E)$ by transforming the abscissa from photon energy to real photo-electron momentum $k = \sqrt{2m(E - E_0)/\hbar}$ where $E_0$ is the energy of the edge step. Historically, the conversion of $\mu(E)$ to $\chi(k)$ has been the “easy” part of XAFS analysis, in which the absorption data was procedurally reduced without much regard to the physical systems being measured. This somewhat brutal treatment of the absorption data is still done here, and $\chi(k)$ is regarded as an intermediate step between the measured spectrum and the physical understanding. The advantage of this approach is that all of the local structural information in the absorption measurement is contained in $\chi(k)$, and yet essentially no assumptions have been made about the physical model for the system.

The $\mu(E)$ term in Eq. (4.8) is the total absorption from a single deep core level, and the $\mu_0(E)$ term is the absorption due to the same deep core level for an isolated absorbing atom in the solid. This will differ at least some from the absorption for a truly isolated atom (in vacuo) because of the overlap of electron orbitals of neighboring atoms. It should still be a smoothly varying function of energy, so that it can be well-separated from the oscillatory part of the total absorption, but its actual form is poorly known. Current state-of-the-art calculations of $\mu_0(E)$ [85] can still only get qualitative agreement with experiments, and their use in analysis is still being explored [85, 86], and is not yet reliable.
The problems for converting measured absorption to $\chi(k)$ are then 1) to get the contribution to the total absorption from only the one edge of interest, and 2) to isolate the atomic absorption from the total absorption for that edge. These two problems are usually solved separately. I will address them separately here, although they are not completely independent in the program AUTOBK, as described below.

### 4.3.1 Pre-Edge Background and Normalization

The first step in converting the total absorption to $\chi(k)$ is to isolate the absorption due to the edge of interest from the absorption due to all other excitations. If the contributions from other edges are not removed they will not affect the fine-structure (the numerator in Eq. (4.8)), so that it doesn’t matter much how well the absorption from the other edges is known. But the normalization of $\chi(k)$ does depend on successfully eliminating the absorption from the other edges. This problem of normalization is compounded by any energy-dependence in the absorption measurement (through energy-dependence in the detector response) and an energy-dependence of the absorption coefficient for the bare atom. Any neglected amplitude corrections of the forms discussed in section 3.1.6 would also complicate the normalization by an energy-dependent $\mu_0(E)$.

Because of these energy-dependent complications, and because the true $\mu_0(E)$ is not well-known, it is common practice to replace the energy-dependent normalization by a constant normalization, independent of energy. This is certainly an approximation of Eq. (4.8), and may not be appropriate for absorption by low Z atoms. But for high Z materials, as analyzed here, the approximation of a constant normalization factor can easily be corrected (as described in section 3.1.6) so as to not give a serious error in the results. The dominant uncertainties in XAFS analysis are in the amplitude terms, and a constant normalization factor is only one among many approximations to be blamed. (The other amplitude terms that cause concern in XAFS analysis are the measurement errors described in chapter 3, and any energy-dependence of the $S_0^2$). The result is that we have a modification of Eq. (4.8) of the form

$$\chi(E) = \frac{\mu(E) - \mu_0(E)}{\Delta \mu_0(E_0)}, \quad (4.9)$$

This constant value of $\Delta \mu_0(E_0)$ used for normalizing the XAFS is determined by taking the difference in the extrapolation of smooth functional fits to the pre-edge total absorption $\mu(E)$ and post-edge background absorption $\mu_0(E)$ at the edge energy,
\( E_0 \), so that
\[
\Delta \mu_0(E_0) = \mu_0^+(E_0) - \mu^-(E_0).
\]

The measured absorption below the edge step (the so-called pre-edge region) is fit to a straight line over an energy region between \([E_0 - 200, E_0 - 50]\) eV. This fitted line is then extrapolated to \( E_0 \), to give \( \mu^-(E_0) \). The background spline (the method for finding this spline is discussed in the next section) is fit to a quadratic polynomial in \( E \) over an energy region between \([E_0 + 100, E_0 + 300]\) eV. This fitted polynomial is then extrapolated to \( E_0 \), to give \( \mu_0^+(E_0) \). The normalization constant is simply the difference of these two extrapolations. Values given here for the ranges are typical, but can be adjusted so as to use reliable data ranges, avoiding glitches and sharp near-edge absorption features.

### 4.3.2 Post-Edge Background

The bigger challenge in converting a measured absorption spectra into \( \chi(k) \) is to isolate the oscillatory part of the spectra from the “smooth” part of the spectra, \( \mu_0(E) \) (usually referred to as “the background”). The principle difficulty is in deciding what is smooth variation due to the absorbing atom, and what is oscillatory structure due to the neighboring atoms. The confusion gets worse near the absorption edge, where atomic contributions can be of comparable amplitude and frequency to the XAFS oscillations. Theoretical predictions for the central atom absorption mentioned above [85] do suggest substantial variations well past the absorption edge, though they are not yet reliable enough to use in data analysis. Progress in theoretical calculations of \( \mu_0(E) \) should eventually be of practical use in XAFS analysis.

Unfortunately, only ad hoc approaches to background removal are currently useful in practice. There has been considerable work on developing practical methods of background removal in recent years [87, 39, 71, 86]. All of these techniques rely on defining what in the measured absorption spectra is not the background, with each using a slightly different definition. All of these approaches have been used successfully and deserve further exploration. For the sake of brevity, I will present here only the method used in this study, which amounts to the status of the algorithms in the computer program AUTOBK as of the date of this thesis. The AUTOBK background removal method described here is fairly general in its approach, and has been used successfully on a wide variety of XAFS problems [71, 80].

The main principle of AUTOBK is to apply the ideas of the information content of an XAFS signal (as in section 4.1) to the separation of the background and XAFS. The
usual assumption for deciding whether a background removal is reliable is that the background is more smoothly varying than the XAFS oscillations. This qualitative criteria can be codified and, with the aid of information theory, made quantitative.

The background \( \mu_0(E) \) in AUTOBK is defined to be a smooth function of energy that, when converted to \( R \)-space, will contain only principle frequencies below some \( R \)-value, \( R_{bkg} \). A piecewise polynomial function, or spline, is used as the smooth background function. Splines are convenient to use because they can be made arbitrarily flexible and are easy to calculate. The splines used in AUTOBK are fourth-order b-splines [88], which are essentially equivalent to regular cubic splines [84], with slightly generalized endpoint conditions. The freedom of a spline is controlled by the points along the \( x \)-axis at which the polynomial pieces are joined, and where they are allowed to be discontinuous in their highest non-trivial derivative. Thus a fourth-order spline is continuous in its value and first two derivatives over its entire range, and is discontinuous in its third derivative at finite number of points, commonly called “breakpoints” or “knots”.

Following the discussion of section 4.1, the knots of the spline are evenly distributed in \( k \)-space (and so quadratically spaced in energy), and there are \( N_{bkg} \) of them, given by

\[
N_{bkg} = \frac{2R_{bkg} \Delta k}{\pi} + 1, \tag{4.11}
\]

where \( \Delta k \) is the \( k \)-range of the data, and \( R_{bkg} \) is some small distance (usually \( \sim 1 \text{ Å} \) or half the distance to the near-neighbor) below which all information in the measured absorption is used as the non-structural background. For a typical XAFS experiment, with \( R_{bkg} = 1.0 \text{ Å} \) and \( \Delta k = 12.5 \text{ Å}^{-1} \), there can be up to 9 knots in the spline. Eq. (4.11) includes the “+1” term to reflect the constant term in the Fourier series because the background spline will contain information starting at 0 Å.

While the energy-values of the \( N_{bkg} \) knots are fixed, the ordinate values for the knots of the spline (through which the spline must pass, and at which the first two derivatives must be continuous) are allowed to take any value — this is what makes the spline adjustable. In this sense the background function \( \mu_0(E) \) depends solely on the set of \( N_{bkg} \) ordinate values of the spline, which I’ll denote as the \( N_{bkg} \)-dimension vector \( \mathbf{y} \), so that \( \mu_0(E) = \mu_0(E, \mathbf{y}) \). The criterion for choosing the best background spline is a mathematical expression of the “smoothness” argument for \( \mu_0 \), and is to find the \( N_{bkg} \) components of \( \mathbf{y} \) which will minimize the non-structural components of the resulting \( \hat{\chi}(R) \) below \( R_{bkg} \).
The function to minimize in the background removal is then

\[
f(R, y) = \text{FT} \left[ \frac{\mu(k) - \mu_0(k, y)}{\Delta \mu_0(k = 0)} - \chi_{\text{model}}(k) \right] \quad R \leq R_{\text{bkg}}, \quad (4.12)
\]

The minimization of \( f(R, y) \) is done in the least-squares sense, using the Levenberg-Marquardt algorithm for non-linear least-squares minimization of a function [84, 89, 90]. Since \( f(R, y) \) is complex, both its real and imaginary components must be minimized. And since \( f(R, y) \) is minimized only over the region below \( R_{\text{bkg}} \), there is no danger of mistakenly removing “real” data above \( R_{\text{bkg}} \), even though \( N_{\text{bkg}} \) may seem like a fairly large number of knots. With this formalism, only low-\( R \) information in the absorption data is used to give the spline, but all of it is used.

To make sure that only the non-structural components of \( \mu(E) \) are removed in the background removal, the expected leakage from the first shell XAFS should be included in the determination of the background function, as is indicated by the term \( \chi_{\text{model}}(k) \) in Eq. (4.12). Since only the spectral leakage (due to the finite Fourier transform range) of this function in the low-\( R \) region will be used, this model needs to be only a rough estimate of the first shell XAFS. Usually a rough calculation from FEFF of the expected first shell XAFS spectra, without any optimization of structural parameters, can be used effectively to prevent the background spline from removing too much of the low-\( R \) components of the absorption spectra.

Figure 4.1 shows an example of the results of AUTOBK. Measured absorption data for pure Cu at 10K is shown along with the background determined using the techniques outlined above. The one free parameter chosen to control the number of knots in the background spline was \( R_{\text{bkg}} = 1.20 \) Å. The Fourier transform used for the background removal had no Fourier transform window smoothing (so that \( dk = 0 \)), and a \( k \)-weighting parameter of \( w = 0 \). With a \( k \)-range for the data extending from \( 0.05 \) Å\(^{-1} \) to \( 25.0 \) Å\(^{-1} \), there were 19 knots used in the spline. A model \( \chi(k) \) from FEFF based on the lattice constant of pure Cu was used to model the leakage of the first shell into the region below \( R = R_{\text{bkg}} \). Figure 4.2 and show the resulting \( k \)-weighted XAFS from this background removal in \( k \)-space. Figure 4.3 shows the \( \tilde{\chi}(R) \) resulting from this background removal, created by Fourier transforming the data shown in Fig. 4.2 with Fourier transform parameters more appropriate for the displaying of XAFS data. This pure Cu data will be used later in this chapter for examples of fitting and error analysis.
Figure 4.1: Measured absorption coefficient (solid) and background from AUTOBK (dashed) for pure Cu at 10K. The background spline used the parameter $R_{\text{bkg}} = 1.2 \text{ Å}$. 19 knots were used to give the smooth background function.

Figure 4.2: Resulting $k^2 \chi(k)$ for pure Cu at 10K after removal of the background determined by AUTOBK.
Figure 4.3: Resulting $\tilde{\chi}(R)$ for pure Cu at 10K after removal of the background determined by AUTOBK. Fourier transform parameters used to make this figure were $w = 2$, and a Hanning window with $k_{\text{min}} = 1.5 \text{ Å}^{-1}$, $k_{\text{max}} = 20.5 \text{ Å}^{-1}$, and $dk = 1.0 \text{ Å}^{-1}$. Both the magnitude (dashed) and imaginary part (solid) of $\tilde{\chi}(R)$ are shown.
4.4 Theoretical Standards in XAFS Analysis

Once $\chi(k)$ has been isolated from the measured absorption spectra, the XAFS equation can be used to extract the atomic species, coordination number, and moments of the partial pair distribution function for the neighboring atoms. Although there are a few different approaches to determining these desired quantities, all of them rely on some prior knowledge of the scattering amplitudes and phase-shifts ($f$ and $\delta$ in Eq. (2.53)) of the system in order to get information about the local structure of the system. The assumptions needed to get these scattering terms are usually fairly modest (including the atomic species of absorbing and scattering atoms and a rough idea of the interatomic spacing) and, in many situations, the results of the local structure are insensitive to the details of the scattering factors. The extent to which the scattering terms influence the results for the local structure has been extensively discussed in the XAFS literature [30].

Historically, there have been two methods used to estimate the standard scattering pieces, $f$ and $\delta$, in the XAFS equation. One way is to use estimates derived from analysis of experimentally measured XAFS spectra of systems for which the atomic distribution around the central atom is known. The other way is to use theoretical calculations of $f$ and $\delta$ from first principles. Both techniques have been used successfully in XAFS analysis, with much of the reliable XAFS analysis done to date using experimental standards. However, due to advancements in theoretical calculations [54, 91, 73] and the difficulties of using experimental standards for analysis beyond the first shell, the use of theoretical standards has become the preferred analysis method in recent years.

Since the analysis in this study is of the temperature dependence of near-neighbors in fcc metals, experimental standards could be used in principle, with the lowest temperature measurement being used as a standard. A fairly simple analysis procedure called the "log-ratio method" [30] could then be used, in which a spectrum from an unknown structure is compared directly to one from a known structure to give the differences in structural parameters. This technique requires XAFS for an isolated "shell" and so is essentially impossible for anything except near-neighbor bonds. Still, it could be used in this work to give the change in structural parameters as a function of temperature. This was, in fact, done on some of the systems in this study. Under favorable conditions, the log-ratio procedure could be coerced to give reasonable results. But even for the first shell it was often found to be too unreliable for this study, tending to give large and spurious systematic errors. This work could include
examples of unreliable results from log-ratio analysis for first-shell XAFS of fcc metals. A careful analysis of mocked-up data (using FEFF calculations with values for $\sigma^2$ put in by hand, for example) using the log-ratio method can easily be shown to give systematic errors in the temperature dependence of XAFS amplitude factors. Analysis of the same data using theoretical standards from FEFF showed no such problem.

Some special considerations must be taken when comparing experimental and theoretical XAFS spectra that can often be neglected when comparing two different experimental spectra. These include difficulties in eliminating weak energy dependencies of the measured absorption coefficient, as mentioned in chapter 3, and Refs. [92, 30]). The basic result from these considerations is to be sure that the experimental $\chi(k)$ be as free as possible from any non-linearities (such as detector saturation or glitches), and that all appropriate corrections (especially for data taken in fluorescence) to the data be made before the analysis is begun. These non-linearities and corrections are more important when using theoretical standards than when using experimental standards because many distortions will cancel (at least to first order) between experimental unknown and experimental standard. The program ATOMS [69] will help calculate most of these corrections.

A second important consideration when using theoretical standards is that the calculations from FEFF are not perfect, especially in the assumptions of the muffin-tin and self-energy approximations of the atomic scattering potentials, and can be noticeably inadequate for some systems. This does not appear to be the case for the simple fcc metals analyzed in this work, but can be quite significant for insulating materials with different atomic species and chemical state. It has been seen, for example, that significant changes in energy-reference are needed to get good and reliable fits for materials in which charge-transfer is known to be important [93, 79]. It might be best to avoid directly comparing experimental XAFS data to model calculations in the most difficult cases, and to compare results of fits for two different experimental data sets instead. Such prudence seems unwarranted for simple close-packed metals, where FEFF is clearly quite good, and muffin-tin approximation seems reasonable. The FEFF results, which have been shown to be quite reliable in XAFS analysis for a variety of systems by many others [77, 78, 79, 94, 76, 58], will be used for close-packed metals in this study without much additional critical analysis of their suitability for XAFS analysis.
4.5 **FEFFIT: Using FEFF in R-space for XAFS Data Analysis**

The analysis program FEFFIT is now described. This program gives an analytic method of determining the local structure around an atom by adjusting model calculations until they best fit XAFS data. The fitting model used in FEFFIT consists of theoretical XAFS spectra from FEFF and parameterized modifications to these that reflect a physical change in the system away from the ideal structure used to make the FEFF calculation. FEFFIT does the fitting of model to data in $R$-space. The discussion presented here is an expansion of the appendix of Ref. [72].

Although it is not necessary to explain the FEFF calculation to create a model XAFS spectra, I will briefly explain the form of the results of the FEFF calculation and how to best use these results to analyze experimental XAFS data. Topics discussed in this section will be the general form of the FEFF solution, the model calculations for XAFS in FEFFIT, and the use of generalized coordinates in the fit to build good physical models with meaningful constraints. The next section will continue the explanation of FEFFIT with a discussion of the criterion used for getting the best-fit and the statistical description of the fit, including goodness-of-fit parameters and the estimate of uncertainties in the fitted variables. The Cu data shown in previous sections will then be analyzed to finish the chapter with a concrete example of fitting and error analysis. Later chapters will, of course, give many more examples of the use of FEFFIT. This work is not meant as a user’s guide to FEFFIT, and more complete program documentation [95] (giving information on how to implement in practice the concepts discussed here) is available upon request [96].

4.5.1 **The Model $\chi(k)$ to Compare to Data**

As mentioned in the introduction to this chapter, FEFF has automated the calculation of multiple-scattering as well as single-scattering contributions to the XAFS. This development allows the analysis of multiple-scattering XAFS data and is one of the most important theoretical advances to the XAFS field made in recent years. In order to best accommodate this ability, FEFFIT uses the FEFF calculations for individual scattering paths, and a sum-over-scattering-paths formalism for the total XAFS, with no conceptual distinction between single- and multiple-scattering paths. The full

---

2 More recent versions of FEFFIT also allow fitting in original or back-transformed $k$-space as well as $R$-space. Though not completely tested at the time of writing, these options appear to work similarly to $R$-space fitting. All fits in this work were done in $R$-space.
model XAFS can then be written as

$$
\chi_{\text{model}}(k) = \sum_{j}^{N_{\text{paths}}} \text{Im}\{\tilde{\chi}_j(k)\},
$$

(4.13)

where $\tilde{\chi}_j(k)$ is the complex contribution to the XAFS from the $j^{\text{th}}$ scattering path from FEFF, representing the entire set of “nearly equivalent” scattering paths. In this sense, FEFF gives one path for all near-neighbors in an fcc metal at finite temperature, even though each central atom has twelve distinct neighbors, and the XAFS spectra is an average of many slightly different near-neighbor distances due to thermal vibrations of the atoms.

In addition to the explicit dependence on $k$, $\tilde{\chi}_j(k)$ depends on both the scattering and atomic distribution details of the scattering path. We start with the form of the XAFS equation given in Eq. (2.53) for the model $\tilde{\chi}_j(k)$,

$$
\tilde{\chi}_j(k) = \frac{f_j(k)}{k(R_j + \Delta R_j)^2} \exp\left\{-2p''R_j - 2p^2\sigma_j^2 - \frac{2}{35}p^4C_4\right\}
\times \exp\left\{i[2kR_j + \delta_j(k) + 2p(\Delta R_j - \frac{2}{35}p^2C_3)]\right\}.
$$

(4.14)

This form is conceptually convenient and the results from FEFF allow it to be used without much modification in practice. The outputs of FEFF (version 5.0 and higher) used by FEFFIT are the \texttt{feffnnnn.dat} files, each of which contains the scattering pieces of the XAFS for the $j^{\text{th}}$ scattering path. These files contain the values for the half-path-length of the ideal scattering path $R_j$, the number of equivalent paths $N_j$, and the geometrical details of the path, including atomic species and Cartesian coordinates of all atoms in the path. They also contain arrays for several terms used in the XAFS calculation as functions of $k$, the real photo-electron momentum referenced to the Fermi level $E_F$. The FEFF arrays include the effective single scattering amplitude for the path $f_j(k)$ and the total phase shift for the path $\delta_j(k)$ \footnote{FEFF actually breaks up $f_j$ and $\delta_j$ into a few different real components when writing its outputs. The total scattering amplitude $f_j$ and scattering phase-shift $\delta_j$ are both trivially reconstructed from these pieces.}. Two other FEFF arrays give the real part of the complex photo-electron momentum to Re$(p)_j(k)$, and the mean-free-path of the photo-electron $\lambda_j(k)$, which is the inverse of the imaginary part of the complex momentum. Both of these pieces of the complex photo-electron momentum are referenced to $E_0$, the bottom of the conduction band, which typically differs from $E_F$ by a few eV.
The fact that two different measures of photo-electron momentum are used in Eq. (4.14) is actually an artifact of FEFF, which calculates the XAFS as a function of \( p \) and then breaks up \( \hat{\chi} \) into functions of \( k \) when writing its outputs because this purely real expression of the photo-electron momentum is closer to the experimentally measured value. The difference between \( k \) and \( p \) is small for \( k \lesssim 2 \) \( \text{Å}^{-1} \) is small, but \( p \) is the more appropriate momentum to use for the moments of the atomic distribution function, as it is the conjugate to the position operator. For this reason, \( k \) is used in Eq. (4.14) as an index of the four FEFF arrays and where explicitly demanded by the way FEFF breaks up its results (which are the \( \exp(i2kR_a) \) term, and the \( 1/k \) term in the leading amplitude factor), and \( p \) is used everywhere else. The \( p \) used for the model \( \hat{\chi}(k) \) is evaluated as

\[
p = p' + ip'' = \sqrt{\left\{ \text{Re}(p_j(k)) + \frac{i}{\lambda_j(k)} \right\}^2 + i\Delta \Gamma_j \frac{2m}{\hbar^2}},
\]

(4.15)

This gives the FEFF value of \( p \) with the possibility of including an additional (and adjustable) broadening term, \( \Delta \Gamma_j \). This term gives a further Lorentzian broadening term to the model calculation and can be used to account for poor experimental resolution or to correct to the mean-free-path \( \lambda \) given by FEFF.

Because a very simple estimate of the Fermi energy is made by FEFF (using the free electron model), and because it is experimentally difficult to determine the \( E_F \) from an absorption edge, an arbitrary shift in energy origin \( \Delta E_0 \) is necessary to compare experimental data to FEFF calculations. The number of different measures of momentum becomes rather confusing at this point. To clear up this confusion, and because it is often important to know what direction a \( \Delta E_0 \) shift moves the \( k \)-grid, I will give a detailed description of the procedure for determining the \( \chi_j(k) \) function for a scattering path, with special emphasis on the evaluation of the two different momentums used. As mentioned above, both the data \( \chi(k) \) and theoretical model \( \hat{\chi}_j(k) \) are discrete functions of \( k \). The \( k \)-values for the experimental data are unchanged in this procedure, but the \( k \)-values for the theoretical model may be shifted. With the discrete \( k \) written as \( k_i = i \delta k \), where \( i \) is an integer index and \( \delta k \) is the grid size (\( \delta k = 0.05 \) \( \text{Å}^{-1} \) is used in FEFFIT). For each grid point \( i \) (and for each path \( j \)), the following steps are taken to evaluate \( \hat{\chi}_j(k) \):

1. The \( \Delta E_0 \)-shifted value of the photo-electron momentum is evaluated according to \( k_i = \sqrt{(i \delta k)^2 - \Delta E_0 - 2m/\hbar^2} \), where \( \delta k \) is the grid step size. A positive \( \Delta E_0 \) will move the theoretical momentums of the photo-electron towards zero, as if the threshold energy reference of the theory was moved up.
2. This “shifted” momentum \( k_i \) is used to obtain values of the \textsc{feff} arrays \( f_j, \delta_j, \) \( \text{Re}(p)_j \), and \( \lambda_j \). Linear interpolation is used for values of \( k_i \) not explicitly given in these arrays.

3. The “shifted” \( \text{Re}(p)_j \), and \( \lambda_j \) are used to evaluate the real and imaginary parts of the complex momentum, \( p_i \) according to Eq. (4.15).

4. The “shifted” \( k_i, p_i, f_j, \) and \( \delta_j \) are all used to create the “shifted” theoretical \( \chi_j(k_i) \) according to Eq. (4.14).

The total XAFS model \( \chi_{\text{model}}(k_i) \) is then built up from each of the individual paths \( \chi(k_i) \) which result from this procedure (including the shifting of \( k_i \)) according to Eq. (4.13). Finally, \( \chi_{\text{model}}(k_i) \) is compared to the data \( \chi(k) \), as will be further described in section 4.6.

The two energy correction terms described above, \( (\Delta E_0, \Delta \Gamma_j) \), the first four cumulants \( (\Delta R_j, \sigma_j^2, C_3_j, \) and \( C_4_j) \), and the overall scale factor \( (S_0^2 N_j) \) in Eq. (4.14) are all potentially adjustable parameters for each path, and are the only parameters which can be adjusted in the theoretical model for \( \chi_j(k) \). Each path parameter is a single-valued number that does not depend on \( k \). These seven path parameters include the physical measurements desired from an analysis of XAFS data (the four cumulants and \( N_j \)) and the terms that cannot be calculated well enough to rely on in practical analysis (\( S_0^2, \Delta E_0 \) and \( \Delta \Gamma \). In practice, \( \Delta \Gamma \) rarely exceeds of 1.0 eV, and can simply be set to zero for most analysis). Together, these path parameters give a fairly general way to modify the \textsc{feff} calculation for \( \chi_j(k) \), and should be able to model the effect of any physical change in the system on the XAFS and to make up for most short-comings of the \textsc{feff} theory.

The overall amplitude term \( (S_0^2 N_j) \) is a combination of two terms that cannot easily be separated for an individual scattering path. \( S_0^2 \) is almost completely independent of both the scattering path and the external physical conditions of the system. And though it is actually expected to be slightly energy-dependent [56], it is taken to be an energy-independent constant for the analysis here. Furthermore, \( S_0^2 \) should depend only on the central atom and energy level excited, and so should be transferable to any system with the same central atom. In this way, the number of equivalent scattering paths \( N_j \), is not difficult to obtain. For single scattering, of course, \( N_j \) gives the number of atoms in that shell.

Before continuing the discussion of the path parameters, some limitations of the approximations in Eq. (4.14) should now be addressed. The most important of these
is that having a non-zero $\Delta R$ in Eq. (4.14) is valid only insofar that $f$ and $\delta$ depend weakly on $R$. From a practical point of view, this means that if $\Delta R/R \gtrsim 0.05$, it is probably necessary to recalculate the scattering pieces given the newly refined atomic positions. A second limitation is that the cumulant expansion is not guaranteed to converge for all systems. It is, however, such a convenient formalism for a model-independent description of the atomic distribution function that no alternative is given in this analysis procedure. Significant departures from the cumulant expansion are rare (it works well in this study of disordered metals at high-temperatures), and would require model-dependent descriptions of the XAFS equation.

4.5.2 Constraining Variables and Path Parameters

Though mentioned in the previous subsection, the fact that each of the seven path parameters can be adjusted for each path needs special attention. In the first place, there is simply not enough information in an XAFS spectra to allow seven adjustable parameters for each of the 10 or so paths typically used in a full multiple-scattering analysis of data. With a reasonably large data $k$-range of $\Delta k = 15 \text{ Å}^{-1}$, there are roughly 10 independent points in each Å of $R$-space data. As the analysis extends to higher $R$, many more paths are involved (the number of scattering paths grows exponentially with path length [76]) while the amount of independent measurements increases only linearly with $R$. Thus the possibility of independently measuring each structural parameter for all paths is remote past the first few shells. As will be shown later in this chapter, there are 6 important paths at the fourth shell distance in an fcc metal. As this shell extends over only $\sim 1 \text{ Å}$, independent measurements of each of the 6 $\sigma^2$ and $\Delta R$ parameters are not possible, and even reliable measurements of a few of them would be difficult. Secondly, some of the path parameters need to be adjusted only in extreme cases — $\Delta \Gamma$ ends up being significantly different from zero only if there is unusually poor experimental resolution, and $C_4$ is non-zero only in cases of very high disorder. Furthermore some of the path parameters will have the same value for different scattering paths — $S_0^2$ and $\Delta \Gamma$ should be the same for different paths in a given material, and $\Delta E_0$ will often be the same for all paths 4. Finally, even if some of the path parameters take a different value for each path, some

---

4 The fact that $\Delta E_0$ is sometimes found to differ significantly for different scattering paths [93, 79] probably reflects a sensitivity of the measured XAFS to chemical bonding that directionally affects the states available to the photo-electron, and the momentum associated with these states that is not completely accounted for in the FEFF calculations.
of these values may be related to one another, so that it is desirable to relate different path parameters to one another.

The necessity and desirability of putting constraints and interdependencies on the path parameters means that explicit instructions be given for how each path parameter is to be adjusted in the fit. This is accomplished in FEFFIT by writing simple mathematical expressions (algebraic and trigonometric functions of real numbers) for each of the path parameters in terms of "generalized variables" specifically chosen and defined for a given problem. As for the background removal problem, we label the set of \( N_{\text{vary}} \) variables chosen for the problem with an \( N_{\text{vary}} \)-dimension vector \( \mathbf{x} \). Unlike for background removal, neither the physical meaning nor the mathematical effect of the variables is specified at this point — this is left as part of the fitting model, and can be different for each analysis.

Formally, we write each of the 7 path parameters (again, these are the terms \( \Delta E_0, \Delta \Gamma, \left(S_0^2, N\right), \Delta R, \sigma^2, C_3 \), and \( C_4 \) in Eq. (4.14)) for the \( j \)th path as functions of the variables \( \mathbf{x} \), so that we have \( \Delta R_j = \Delta R_j(\mathbf{x}) \), \( \sigma_j^2 = \sigma_j^2(\mathbf{x}) \), etc. The actual functional form can be as simple as a \( \Delta R_2 = 0.01 \) to set the change in path length for the second path to the constant value 0.01 (so that \( \Delta R_2 \) actually depends on none of the variables). Or it can be as complex as \( \sigma^2 = \max(0.10, \sqrt{\text{therm}^2 + \text{static}^2}) \), which uses intrinsic functions \( \sqrt{\text{therm}^2 + \text{static}^2} \) in the FORTRAN sense (for square-root and the maximum of two numbers, respectively) to set the value of \( \sigma^2 \) for the first path to be a combination of the two terms \text{therm} and \text{static}, with the added restriction that the value not be greater than 0.10. The terms \text{therm} and \text{static} are defined for the particular problem, and can be used in other path parameter equations. Either (or neither or both) of \text{therm} and \text{static} can be explicitly varied in the fit. They could also be fixed (as was \( \Delta R_2 \) above), or could depend on a combination of other variables. The formalism outlined here has been found to give enough freedom to allow a physically meaningful, minimal set of "generalized variables" to be used for a wide variety of XAFS problems. Despite the formal-sounding nature, it is transparent and convenient enough to use in practice (though I may be biased).

An example of a simple constraint for path parameters of different paths is to require \( \Delta E_0 \) to be equal for all scattering paths in a problem, even if that value is allowed to vary in the fit. A second example involves expansion of a crystal (if, for example, the FEFF calculation started with a slightly incorrect lattice constant). All path lengths would need to be adjusted by a different amount, but there is really only one physical parameter (the expansion coefficient) that can used to write the change
in half-path-length $\Delta R$ for each of the paths. As a third example, the number of near-neighbors of two different atom species in an fcc alloy could be determined by requiring the values of $N$ for the two scattering paths for the two different backscatters to add to 12, thereby using only one free physical parameter for the XAFS amplitude factor of two paths. Such a tactic will be used to measure the impurity concentration in the dilute alloys in chapters 8 and 9. A fourth example of a physically meaningful constraint would be to have all mean-square displacements $\sigma^2$ for a system be described by a single Debye Temperature. This reduces many physical quantities (one for each path) to only one adjustable parameter.

More complex sets of constraints can be used when more than one data set are fit simultaneously. The importance of simultaneously using more than one data set is greatly enhanced by a being able to write general equations of constraints for path parameters with variables that extend over different data sets. For example, the above discussion of a single expansion coefficient giving change in path-length for many multiple-scattering paths can be enlarged to include using many different temperatures and a single thermal expansion coefficient to give all the $\Delta R$'s for all paths and all temperatures. The temperature-dependence of $\sigma^2$ (or any or all cumulants) can be similarly constrained in this way to a reduced set of generalized thermal parameters. Different data sets that are to be fitted simultaneously don't necessarily have to differ in temperature (though that is the only type of difference presented in this work). It is possible, for example, to simultaneously fit the XAFS from two different central atoms, and to constrain distances and mean-square-displacements to be consistent for the two different measurements.

An even more complicated set of path parameter constraints, which models the temperature dependence of the four cumulants for the near-neighbor bond in terms of a set of three anharmonic force constants at any temperature, will be used later in this study. These constraints reduce four path parameters for many different temperatures to a total of only three variables. This topic will be further developed in chapter 5 where the equations of constraint will be presented, and in chapters 6, 8, and 9, where the constraints will be used with real XAFS data to measure the three anharmonic force constants. Appendix A lists an actual set of constraints used in chapter 6 to fit the temperature-dependent Ag data.
4.5.3 Integrating AUTOBK and FEFFIT

The claim was made in section 4.3.2 that the structural and non-structural portions of XAFS spectra could be well-separated in $R$-space with a sharp cut-off, or something quite like it. This gives the single parameter $R_{\text{bkg}}$ below which a spectrum is all background and above which it is all structural signal. Such a bold statement needs more justification, and the interdependence of the signal below and above $R_{\text{bkg}}$ should be further explored. In addition, the Fourier transform parameters for background removal and analysis are often different (the $k$-weighing parameter, for example, is usually smaller for background removal than for structural analysis). The upshot of all this is that it would convenient, interesting, and useful if the background and structural parameters could be determined together.

The solution adopted in FEFFIT is to assume that AUTOBK gives a reasonably good background, but that changes in the background spline should be allowed so as to reduce the low-$R$ components of the fit in $R$-space. So as to make this an option that can be easily turned on and off, FEFFIT uses the $\chi(k)$ form of the data whether or not it includes a background function. The full model XAFS then changes from that given in Eq. (4.13) to

$$
\chi_{\text{model}}(k) = \Delta\mu_0(k; x_{\text{bkg}}) + \sum_{j}^{N_{\text{pars}}} \text{Im}\{\tilde{\chi}_j(k; x_{\text{struct}})\}.
\tag{4.16}
$$

The physical and mathematical meaning of the set of structural variables $x_{\text{struct}}$ are still chosen for an individual problem. As in AUTOBK, the set of background variables $x_{\text{bkg}}$ are the ordinate values of the spline, and there are $N_{\text{bkg}}$ of them, as given by Eq. (4.11) with $R_{\text{bkg}} = R_{\text{min}}$, the low-$R$ end of the structural region. The spline is evaluated in $k$-space directly, and $\Delta\mu_0(k, x_{\text{bkg}})$ which is the change in the background function for the AUTOBK spline is expected to be small. The fit $R$-region is now increased from $[R_{\text{min}}, R_{\text{max}}]$ to $[0, R_{\text{max}}]$ with the region $[0, R_{\text{min}}]$ being devoted to background and the region $[R_{\text{min}}, R_{\text{max}}]$ being devoted to structure.

While this gives a pleasing combination of the AUTOBK and FEFFIT approaches to XAFS analysis (and also tends to make nice pictures and slow the processor down), there is also a real analytic benefit from this approach. The question of the interdependence of background and structural parameters can now be directly and quantitatively investigated. As discussed in the next section, FEFFIT determines the uncertainties in the fitted variables and the correlation between pairs of variables. This statistical analysis is done for all variables so that when $\Delta\mu_0(k, x_{\text{bkg}})$ is included in the fit the correlation of the background variables and the structural variables are
determined. The relative size of the uncertainties and best-fit values for the background variables tells how badly the background needed fixing. And the correlations between background and structural variables tells how much the background and structural regions influence one another.

4.6 Goodness of Fit and Estimation of Uncertainties

With all the scattering information from path (in the form of the four arrays from FEFF), and the set of path parameters written in terms of the desired variables \( \mathbf{x} \) in the problem, we can turn our attention to getting the best values for the set of variables. After that, we can evaluate statistical quantities of the fit, including measures of the goodness-of-fit and estimates of the uncertainties in the fitted variables. The procedures for these steps used in FEFFIT come from standard non-linear least-squares minimization and standard error analysis techniques [89, 84, 90]. Since these are such an important part of the analysis, I will discuss them in some detail.

4.6.1 Fitting Criteria and Measures of the Goodness of Fit

As with the discussion above for background removal, we need to specify a function to be minimized in the least-squares sense. As with AUTOBK, FEFFIT evaluates this function in \( R \)-space. In analogy to the form of Eq. (4.12) for the function to minimize for background removal, the function to minimize for the analysis of the \( \chi(k) \) data is

\[
f(R, \mathbf{x}) = \text{FT}\left[ \chi_{\text{data}}(k) - \chi_{\text{model}}(k, \mathbf{x}) \right], \quad R_{\text{min}} \leq R \leq R_{\text{max}},
\]

where I've explicitly written the dependence of the model XAFS on the user-chosen variables \( \mathbf{x} \). As for AUTOBK, the Levenberg-Marquardt method of non-linear least-squares minimization [84, 89, 90] is used to find the best set of the variables \( \mathbf{x} \) as those that minimize the sum of squares of the components of \( f(R, \mathbf{x}) \). Real and imaginary parts of \( f(R) \) of Eq. (4.17) are used in the minimization with equal weight\(^5\). The finite \( R \)-range used for \( f(R, \mathbf{x}) \) gives some spatial filtering of the data, so that high-\( R \) regions, where multiple- and single-scattering paths become too numerous and distinct to be accurately modeled, can be ignored.

---

\(^5\)This gives a better fit to the phase than the magnitude, but that this may actually be desirable because XAFS measurements themselves are usually more accurate in the phase than in the amplitude, as discussed in chapter 3. It is not clear whether there is actually more information about the phase than the amplitude.
After a best-fit is found, it is often desirable to have some measure of the goodness-of-fit. The most common measure of goodness-of-fit is the $\chi^2$ statistic, which reports a weighted sum of squares of the difference of data and model, and is therefore quite closely related to the function $f(R,x)$ minimized in the fit. The standard definition of $\chi^2$,

$$\chi^2 = \sum_{i=1}^{N} \left( \frac{f_i}{\epsilon_i} \right)^2,$$

requires that both $N$, the number of evaluations of $f$, the function to minimize, and $\epsilon_i$, the uncertainties in this function, be known. The index $i$ here indicates the different $R$ values. We want the fit to extend from $R_{\text{min}}$ to $R_{\text{max}}$, so we have $f_i = f(R_{\text{min}})$ and $f_N = f(R_{\text{max}})$. With an $R$-space grid spacing of $\delta R$, $N = (R_{\text{max}} - R_{\text{min}})/\delta R$. The value of $\delta R$ is arbitrary and depends on the array size of the fast Fourier transform (see section 4.2). A better number to use here would be the spacing between independent points, so that $N$ would be $N_{\text{idp}}$, the number of independent points in the fit region. But the fast Fourier transform has a fixed maximum frequency (which is $2\pi/\delta k \approx 31.416\,\text{Å}$ when $\delta k = 0.05\,\text{Å}^{-1}$), and is much faster for when the size of the array that extends up to this maximum frequency is a power of two. Since the spacing between independent points rarely gives a power of two for this maximum frequency, $N$ (which is the number of evaluated points in the $R$-region of the fit) is chosen so that the number of data points in the whole frequency range is a power of two, and so that $N$ is “slightly” larger than $N_{\text{idp}}$. $N$ is rarely set to more than twice $N_{\text{idp}}$. With $N$ evaluated points over the $R$-range which has $N_{\text{idp}}$ points, we have

$$\chi^2 = \frac{N_{\text{idp}}}{N} \sum_{i=1}^{N} \left( \frac{f(R_i)}{\epsilon_i} \right)^2.$$

Unlike for background removal, the fit range does not include the origin, so $N_{\text{idp}}$ is strictly an even number, with pairs of independent points spaced at intervals of $\pi/2\Delta k$. As pointed out by Stern [82], the fit range in $R$-space can be chosen so that the endpoints of the fit ranges are strictly on the grid spacing of independent points, so that the number of independent points in the fit range (including endpoints) is $N_{\text{idp}} = 2 + 2\Delta R \Delta k/\pi$. Due to the finite $R$-grid imposed by the discrete Fourier transform, the exact alignment of the Fourier transform $R$-grid with the $R$-spacing of independent points is unlikely. Since the endpoints of the fit range cannot be guaranteed to reside on independent points in the data, the number of independent points completely inside the fit range is used instead.

The term $\epsilon_i$, which gives the uncertainties in the measurement of $\tilde{\chi}(R_i)$, is usually not known ahead of time, and is not trivial to determine for a given spectra.
Furthermore, the statistical analysis of the data based on this $\chi^2$ statistic turns out to depend rather strongly on having $\chi^2$ scaled correctly, so that a good estimate of $\epsilon_i$ is needed to get reliable measures of the goodness-of-fit and uncertainties in the measured variables. Although it is often possible to estimate the uncertainty in the measurements of $\mu(E)$, the uncertainties in $\hat{\chi}(R)$ are quite a different matter (they include $k$-weighting, for example). It turns out that the random (or “white noise”) portion of $\epsilon$ can be estimated, as will be discussed below. But before presenting that discussion, let’s finish the definition of $\chi^2$ by further assuming that $\epsilon$ does not vary strongly with $R$, so that we have a single value of $\epsilon$ that can be pulled out of the sum\(^6\), giving

$$
\chi^2 = \frac{N_{\text{dof}}}{N} \sum_{i=1}^{N} \left( \hat{\chi}_{\text{data}}(R_i) - \hat{\chi}_{\text{model}}(R_i; \mathbf{x}) \right)^2. \tag{4.20}
$$

The random component of $\epsilon$ can be estimated fairly simply. To do this we rely on the decay of the XAFS signal from the mean-free-path and Debye-Waller factors at high-$R$, and the observation that the random noise will have roughly constant amplitude in $R$-space (i.e. it will be “white”). The XAFS signal will then be indistinguishable from the noise at sufficiently large $R$. With a typical mean-free-path of $10 - 20$ Å, we empirically take $\epsilon$ to be the rms average of the $\hat{\chi}_{\text{data}}(R)$ between 15 and 25 Å as the rms value of the random noise. This value certainly ignores some systematic errors in $\hat{\chi}_{\text{data}}(R)$, which may very well dominate the measurement uncertainty. These might include a wide variety of systematic errors in the measurement of $\mu(E)$ which filter through to $\hat{\chi}_{\text{data}}(R)$, errors in background removal, and errors due to the linear interpolation of $\chi_{\text{data}}(k)$ from the original data (in tests on mocked up data with known amount random noise added, both background removal and linear interpolation errors seem to be quite small). This simple evaluation of $\epsilon$ gives a good estimate of the random noise in $\hat{\chi}(R)$ without any prior knowledge of the system (so that a computer program can estimate a value for $\epsilon$ given a single spectrum of $\chi(k)$ data and no other information). It also takes into account all scaling effects from the Fourier transform. If the noise is dominated by random fluctuations (i.e., if it is “white”), the relation between the noise in $k$- and $R$-space is linear [97], and is given by

$$
\frac{\epsilon_k}{\epsilon_R} = \sqrt{\frac{\pi (2w + 1)}{\delta k (k_{\max}^{2w+1} - k_{\min}^{2w+1})}}, \tag{4.21}
$$

\(^6\) The random component of $\epsilon$ does not vary with $R$, and we have little reason to expect that systematic errors in the data $\hat{\chi}(R)$ would vary strongly with $R$. 
where, as in section 4.2, $\delta k$ is the $k$-space grid, $k_{\text{min}}$ and $k_{\text{max}}$ are the Fourier transform $k$-range, and $w$ is the $k$-weighting parameter.

The first use of the $\chi^2$ statistic is as a direct measure of the goodness-of-fit. In this way, a fit is usually deemed “good” if $\chi^2 \approx \nu = N_{\text{dp}} - N_{\text{varys}}$, which is to say that the model and data differ by $\sim \epsilon$ (the noise in the data) at each of the points in the fit, and that the each of the $N_{\text{varys}}$ variables improve the fit at one of the points. A second figure-of-merit (called reduced chi-square) can be defined as $\chi_r^2 = \chi^2 / \nu$ at this point. The above criterion for a good fit is then $\chi_r^2 \approx 1$. $\chi_r^2$ can also be used to compare two different fits: a fit with a lower $\chi_r^2$ being considered the better of two fits. This comparison doesn’t depend strongly on the value $\epsilon$ (as long both fits aren’t accidently too good), and is the preferred way to compare fits with different numbers of variables. In this way, $\chi_r^2$ can be used to decide if a variable is needed in a fit — if an added variable does not improve $\chi_r^2$, it is not needed. It must be noted, however, that the scale of both $\chi^2$ and $\chi_r^2$ depend on the scale of $\epsilon$.

Although we will return to the $\chi^2$ statistic (and the important consequences of not knowing $\epsilon$) in the next subsection when discussing the estimate of uncertainties in the fitted variables, we now present an alternative figure-of-merit to $\chi^2$ and $\chi_r^2$ for the judging the quality of a fit. Following the statistical analysis used in x-ray diffraction (which has a similar difficulty in estimating the measurement uncertainty), we define an XAFS reliability factor, $\mathcal{R}$, as

$$
\mathcal{R} = \frac{\sum_{i=1}^{N} \left\{ [\text{Re}(f(R_i))]^2 + [\text{Im}(f(R_i))]^2 \right\}}{\sum_{i=1}^{N} \left\{ [\text{Re}(\tilde{f}_{\text{data}}(R_i))]^2 + [\text{Im}(\tilde{f}_{\text{data}}(R_i))]^2 \right\}}
$$

(4.22)

where the sum is done as for Eq. (4.20). This $\mathcal{R}$-factor is similar to one proposed in Ref. [31] for judging the quality of a fit. It measures the misfit relative to the amplitude of the data, and is used to give some confidence in the quality of the fit. It is proportional to $\chi^2$, and is a reasonable measure of the goodness-of-fit because it is directly proportional to the quantity actually minimized in the fit. But it avoids the scaling in $\chi^2$ which relies on a good estimate of $\epsilon$. The combination of $\mathcal{R}$ and $\chi^2$, which give the misfit relative to the scale of the data and the misfit relative to the random fluctuations in the data, respectively, is often informative in estimating how much of the misfit is due to a poor model (which would include a poor set of variables and intrinsic problems in the FEFF calculation) and how much is due to poor data. This combination can also give a sense of the size of the systematic errors in the data.
4.6.2 Estimation of Uncertainties for Fitted Variables

Uncertainties for the \( N_{\text{vars}} \) variables in the fit, as well as the correlations between pairs of variables, can be found using standard techniques of statistical analysis \([89, 84]\). The goal of the least-squares minimization is to find the smallest value of \( \chi^2(x) \) in the \( N_{\text{vars}} \)-dimensional space. Of course, the scaling factors don’t matter for this part of the procedure and minimizing \( R \) or \( f(R) \) itself would work just as well. I will continue to consider \( \chi^2(x) \) to be the quantity to be minimized. Near the best solution (\( i.e. x = x_0 \)), \( \chi^2(x) \) will be roughly parabolic in each of its dimensions (it will be at a local minimum in each dimension, so that as a variable is moved away from its best solution in any direction, \( \chi^2 \) will increase.). Most techniques to minimize \( \chi^2(x) \), including the Levenberg-Marquardt method, evaluate the curvature of the \( \chi^2(x) \) surface in order to find the best solution. This curvature give a measure of how sensitive \( \chi^2(x) \) is to changing the value of a particular variable, which exactly the same kind of measure wanted for the uncertainty in the variable — it measures by how much the value of a variable can be changed without significantly changing the value of \( \chi^2(x) \).

The \( N_{\text{vars}} \times N_{\text{vars}} \) curvature matrix (also known as the Jacobian matrix, and containing the second derivatives of \( \chi^2 \) with respect to the variables \( x \)) can be inverted to give the correlation matrix, which gives an estimate of the uncertainties of each the variables as well as the full set of correlations between the variables. Because the curvature matrix is inverted (\( i.e., because the unity matrix is divided by the curvature matrix to give the correlation matrix \)), the uncertainties estimated in this way will increase \( \chi^2 \) by 1, and the scaling difficulties due to not knowing \( \epsilon \) return.

Figure 4.4 shows a crude rendition of a “contour” plot of the \( \chi^2 \) surface for a two-variable problem. At the solution, the variables \( x \) and \( y \) have values \( x_0 \) and \( y_0 \), and \( \chi^2 = \chi_0^2 \). As \( x \) or \( y \) move away from their best-fit solution, \( \chi^2 \) increases. The contours of constant \( \chi^2 \) will be ellipses for two dimensions (and higher order ellipsoids for more than two dimensions) as long as the uncertainties are “normally distributed”. The uncertainty in the value of a variable is the amount by which it can be increased and still have \( \chi^2 \) below some limit. The actual limit is open for debate, but \( \chi_0^2 + 1 \) is one common standard, and the one used in this work. It is also the one that comes from a strict inversion of the curvature matrix. From Fig. 4.4, the uncertainties in \( x \) and \( y \) are \( \Delta x \) and \( \Delta y \), and are those values which demand that \( \chi^2 \) is increased by 1 from its best value. Note that when evaluating the uncertainty in a variable, all the other variables have to be allowed to vary, so the correlation between
Figure 4.4: A contour map of $\chi^2$ as a function of two variables, $x$ and $y$. The uncertainties in the variables ($\Delta x$ and $\Delta y$, respectively) are chosen so as to require that $\chi^2$ is increased by 1 from its best value, $\chi^2_0$. The correlation between the variables $x$ and $y$ is given by $C(x,y) = \cos(\theta_c)$.

Variables can be taken into account. If $y$ were held constant, the uncertainty in $x$ would be considerably smaller than $\Delta x$. The correlation of two variables, $C(x,y)$ is a measure of how much the best-fit value of one of the variables changes in response to changing another variable away from its best-fit value. In Fig. 4.4 the correlation of the variables $x$ and $y$ is given by $C(x,y) = \cos(\theta_c)$, the angle made by the points A, $\chi^2_0$, and B. If the variables were completely uncorrelated, the ellipse would have its major and minor axes parallel to the $x$ and $y$ axes.

Since the estimation of the uncertainties from the inversion of the curvature matrix depends on an accurate value of $\epsilon$, which is generally unavailable, we use a trick to get a reasonable scale for the uncertainties. The idea is to admit that we do not know what $\epsilon$ is, but that we’re confident that the fit is reasonably good. (If the fit was clearly wrong we probably wouldn’t care what the uncertainties in the variables were.) We then can pick the value of $\epsilon$ to give $\chi^2 = \nu$, so that we have a good fit by definition. The estimated uncertainties which increase $\chi^2$ by 1 will then be the values by which variables must be increased to make $\chi^2$ worse by a factor of $1/\nu$. This procedure (which is common in analysis of physical data, even when not admitted) can be applied after a fit is completed by assuming that the errors are
Table 4.1: List of important paths from \textsc{feff} for 10K Cu. Paths are enumerated according to increasing \( R \), and are listed with half-path-length \( R_{\text{eff}} \), degeneracy \( N_{\text{degen}} \), number of scattering sites \( N_{\text{scatt}} \), and Zabinsky curved-wave importance factor \( C_{\text{cw}} \). Paths 10 and 12 are the “focused” multiple scattering paths at twice the near-neighbor distance.

<table>
<thead>
<tr>
<th>Path Index</th>
<th>( R_{\text{eff}} ) (Å)</th>
<th>( N_{\text{degen}} )</th>
<th>( N_{\text{scatt}} )</th>
<th>( C_{\text{cw}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.548</td>
<td>12</td>
<td>1</td>
<td>100.0</td>
</tr>
<tr>
<td>2</td>
<td>3.603</td>
<td>6</td>
<td>1</td>
<td>23.0</td>
</tr>
<tr>
<td>3</td>
<td>3.828</td>
<td>48</td>
<td>2</td>
<td>10.6</td>
</tr>
<tr>
<td>4</td>
<td>4.349</td>
<td>48</td>
<td>2</td>
<td>8.7</td>
</tr>
<tr>
<td>5</td>
<td>4.413</td>
<td>24</td>
<td>1</td>
<td>55.4</td>
</tr>
<tr>
<td>6</td>
<td>4.754</td>
<td>48</td>
<td>2</td>
<td>10.6</td>
</tr>
<tr>
<td>7</td>
<td>4.754</td>
<td>96</td>
<td>2</td>
<td>22.0</td>
</tr>
<tr>
<td>8</td>
<td>5.096</td>
<td>12</td>
<td>1</td>
<td>19.0</td>
</tr>
<tr>
<td>9</td>
<td>5.096</td>
<td>12</td>
<td>2</td>
<td>8.5</td>
</tr>
<tr>
<td>10</td>
<td>5.096</td>
<td>24</td>
<td>2</td>
<td>43.9</td>
</tr>
<tr>
<td>11</td>
<td>5.096</td>
<td>12</td>
<td>3</td>
<td>8.3</td>
</tr>
<tr>
<td>12</td>
<td>5.096</td>
<td>12</td>
<td>3</td>
<td>33.0</td>
</tr>
</tbody>
</table>

normally distributed and then scaling all uncertainties by a factor of \( \sqrt{\chi^2} \). In this way, the reported uncertainties are those which will increase \( \chi^2 \) by an amount equal to \( \chi^2 \). This rescaling will be done for all uncertainties reported in this work.

4.7 An Example Fit; Pure Cu at 10K

As an explicit example of \textsc{fefffit}, I return to the pure Cu data shown in section 4.3.2. A \textsc{feff} calculation for pure Cu is quite easy to do and can be based only on the fcc lattice constant (3.603Å)[98]. The resulting calculation gives 12 paths out to the fourth neighbor distance of 5.10Å (after the Zabinsky importance filters [76] have eliminated scattering paths that give little contribution to the XAFS). At this distance, focused forward multiple-scattering is quite strong. Table 4.1 lists the scattering paths used, their half-path-length, degeneracy, number of scattering sites, and relative Zabinsky curved-wave importance factor.

As a preliminary check of the method, the fit to the 10K Cu data was first done
to only the first "shell". For this analysis, only the first three paths listed in Table 4.1 were used (the second and third path were included because they both leak into the $R$-range of the first path). The Fourier transform parameters used were $w = 2$, $k_{\min} = 1.5 \text{ Å}^{-1}$, $k_{\max} = 18.5 \text{ Å}^{-1}$, and $dk = 1.0 \text{ Å}^{-1}$. The first shell was defined by $R_{\min} = 1.60 \text{ Å}$, and $R_{\max} = 2.75 \text{ Å}$, so that there were 10 independent data points.

Single values of $\Delta E_0$, $S_0^2$ for all three paths were adjusted in the fit, as was a single linear expansion parameter $\alpha$ (so that for each path $j$, $\Delta R_j = \alpha R_{\text{eff}}$). Also varied in the fit was a single Debye Temperature $\theta_D$, which gave the values for $\sigma_j^2$ for the three paths, according to the method described in [99].

The fit results are shown in Fig. 4.5. The random component of the uncertainty in the data was found to be $\epsilon \approx 0.039 \text{ Å}^{-3}$ (since the Fourier transform was done with $k^2$ weighting, both $\tilde{\chi}(R)$ and $\epsilon$ have these units). The goodness-of-fit parameters were $\mathcal{R} = 1.94 \times 10^{-3}$, $\chi^2 = 124.7$, and $\chi^2_{\nu} = 20.8$ (with $\nu = 6$). The best-fit values and estimated uncertainties of the four fitted variables were $S_0^2 = 0.94 \pm 0.03$, $\Delta E_0 = 0.50 \pm 0.46 \text{ eV}$, $\theta_D = 285.4 \pm 16.8 \text{ K}$, and $\alpha = 0.0002 \pm 0.0008$. The variables $S_0^2$ and $\theta_D$ were found to be fairly highly anti-correlated ($C(S_0^2, \theta_D) \approx -0.87$), and the variables $\Delta E_0$ and $\alpha$ were found to be fairly highly correlated ($C(\Delta E_0, \alpha) \approx 0.86$).
Figure 4.6: Pure Cu data at 10K (dashes) and the combination of the twelve scattering paths listed in Table 4.1 (solid). The theory uses the results of the fit to the first shell of Cu, shown in Fig. 4.5 and discussed in the text. No variables were adjusted past the first shell.

The variable $\alpha$ can be taken out of the fit without affecting $\chi^2$ (and, in fact, if $\alpha$ is set to zero so that only three variables are fitted, $\chi^2_\nu$ is reduced without significant changes to the values of the other variables). This means that the measure of the near-neighbor distance from the XAFS analysis using FEFF agrees with the bulk results to within the uncertainties in the XAFS data. The best-fit value and uncertainty for $\alpha$ can be translated to a near-neighbor distance of $R_1 = 2.5484 \pm 0.0020$ Å, while the bulk value (at 10K) is $R_1 = 2.5478$ Å.

If the results for the parameters fit to the first shell are applied to all twelve scattering paths listed in Table 4.1, with no additional fitting done of the parameters, results like those of Fig. 4.6 can be obtained. It can be easily seen that a small number of paths from FEFF will give a very good approximation to the XAFS with a suitably small number of adjustable parameters. The fit to the further shells can easily be done by allowing additional distance and disorder parameters. Simultaneous fits to more than one data set can also be done easily. Such fits, to the temperature dependent XAFS data, are presented in detail in chapters 6, 8, and 9.
Chapter 5

THE STRUCTURAL AND THERMODYNAMIC PROPERTIES OF AN ANHARMONIC POTENTIAL

In this chapter a simple and general model for the near-neighbor radial distribution function, consisting of a one-dimensional anharmonic potential, is presented. The effect on the XAFS for this radial distribution function is developed through the cumulant expansion. The local vibrational contribution to the thermodynamic partition function is found for this anharmonic potential, and thermodynamic quantities for the near-neighbor bonds are derived from it. Since XAFS measurements are of the radial distribution of bond length between probe and near-neighbor backscattering atoms, the disorder in the radial distribution affecting the XAFS signal comes from vibrations in the near-neighbor bond. This gives an description of the solid similar to that of the Einstein model, with all vibrations in the near-neighbor bond occurring at one effective frequency. In this way, all long-wavelength phonons are ignored in the XAFS measurement, in favor of measurements of bond strength. This approach has been shown to successfully explain the thermal disorder in XAFS for many systems [30, 36, 35, 34, 39]. The development here of the XAFS for an anharmonic potential between near-neighbors is based on that of Stern, Līviņš, and Zhang [39], who used the same formalism to describe the thermal disorder in the XAFS of pure Pb up to and even including the melting temperature. The work here extends their approach to include the thermodynamics of this potential. This extension will be used in later chapters to evaluate thermodynamic parameters about interatomic bonds from the analysis of XAFS data.

We parameterize the interatomic forces between near-neighbors in a solid by a one-dimensional anharmonic potential. Putting one of the atoms at the origin (for convenience, the absorbing atom is placed at the origin), such a potential can be expanded in powers of \( r = R - R_0 \), where \( R \) is the instantaneous interatomic distance and \( R_0 \) is the nominal value for this distance, as

\[
\psi(r) = \frac{1}{2}ar^2 + br^3 + cr^4.
\]

This potential gives the mean, central force on the backscattering atom at \( R = \)
$R_0 + r$. It is assumed that $\psi(r)$ does not depend on temperature and that classical thermal excitations are dominant, so that the temperature dependence of the partial pair distribution function is given by $g(r, T) \propto \exp(-\psi(r)/k_B T)$, where $k_B$ is the Boltzmann constant. This $g(r, T)$ gives the temperature-dependent probability of finding an atom at a position $R = R_0 + r$ from the central atom, and is written as

$$g(r, T) = \exp\left[\frac{-\psi(r)}{k_B T}\right]$$

(5.2)

where $\psi(r)$ is the average “effective” interaction for pairs of atoms.

It should be emphasized that $\psi(r)$ is not the usual pair potential used in most theoretical descriptions of condensed systems, which would give the interaction between two isolated atoms as a function of their relative positions. Rather, $\psi(r)$ is the effective pair potential or “potential of mean force” (the term used by Cusack [18], Waseda [16], and March [100], to name this quantity. Cusack gives an excellent discussion of this subtle difference in descriptions of interatomic forces) that measures the net interaction experienced by pairs of atoms. The influence of all other atoms in the solid is implicit in this net interaction that makes up $\psi(r)$, through the thermal averaging over their distributions. These influences include two-body and higher-order interactions between atoms other than the two atoms under consideration. In this way, the consideration of only near-neighbor forces in this work is compensated for.

Converting from $\psi(r)$ to a genuine pair potential is quite difficult in general as these other interactions would need to be accurately accounted for. But such a conversion is probably not even interesting, since, in a practical sense, the $\psi(r)$ here is the better description of the solid than the rather artificial pair potentials. The direct use of the potential of mean force in theoretical descriptions of metals would be quite advantageous at this point. The empirical embedded-atom-method (EAM) [6, 7] for molecular dynamic simulations of materials, which has been quite successful in describing close-packed metals, uses effective pairwise interatomic potentials much like the potentials used for this work. It might be interesting to try to relate the potentials measured here with those used in such simulations, and possibly to

---

1 It is common to cut off calculations of interatomic forces in metals at distances of a few times the near-neighbor distance. The predominance of short-range interactions in metals is usually attributed to the screening abilities of the conduction electrons, but the important result is that pairwise interactions up to a few near-neighbors are usually found to be sufficient to describe metals. This strengthens the case here for using only the effective near-neighbor interactions to describe the thermodynamic properties of fcc metals.
use the XAFS-derived potentials in molecular-dynamics calculations. This would be particularly advantageous for simulations of systems with dilute impurities, for which the usual methods of determining the potentials to use in such simulations can be troublesome and unreliable.

Since the analysis in this work is done for systems at moderately high temperatures, the classical limit will be used for thermodynamic quantities, and the zero-point motion of the atoms will be ignored. The effect of zero-point motion on anharmonicity in systems at low-temperatures has been included in some XAFS analyses [101], but will not be included here, where the discussion is restricted to the classical regime \((T \gtrsim \theta_D/2)\). We will further assume that the anharmonicity is small enough at all temperatures that the cubic and quartic terms \((b \text{ and } c \text{ in Eq. (5.1)})\) can be treated as perturbations of the harmonic potential. (The formal conditions for the anharmonicity to be small enough for the approximations and expansions used below to converge are \(|b| \ll a(a/k_B T)^{1/2}\) and \(|c| \ll (a^2/k_B T)).\)

### 5.1 The Cumulants for a Slightly Anharmonic Potential

The temperature-dependent moments of the pair correlation function \(g(r, T)\) due to thermal motion of the backscattering atom are

\[
\langle r^n \rangle_T = \frac{\int dr \, r^n \, e^{-\beta \psi(r)} \int dr \, e^{-\beta \psi(r)}}{\int dr \, e^{-\beta \psi(r)}}. \tag{5.3}
\]

For XAFS analysis it is more usual to describe thermal disorder with cumulants described in section 2.4 rather than the more usual moments. As described in that section, the cumulants are the coefficients of an expansion in powers of \((2ip)\) of the logarithm of the statistical average \(\langle e^{i2pR} \rangle_T\). They can rather easily be written in terms of the more usual moments of equal and lower order [61]. In this work, the 4 lowest cumulants will be considered to be sufficient to describe \(g(r, T)\), which can be written in term of the lowest four moments as given in Eq. (2.50). To work out the integrals for \(\langle r^n \rangle_T\), we assume that the anharmonicity is small enough so that we can expand the anharmonic portions of the potential in a power series, which gives

\[
e^{-\beta V(r)} = e^{-\frac{1}{2} \beta \sigma^2 \left[1 - \beta (br^3 + cr^4) + \frac{1}{2} \beta^2 (br^3 + cr^4)^2 \\
- \frac{1}{6} \beta^3 (br^3 + cr^4)^3 + \frac{1}{24} \beta^4 (br^3 + cr^4)^4 + \cdots \right]} \tag{5.4}
\]
Where we use $\beta = 1/k_B T$ to simplify the notation. We can then use the integral identity
\[
\int_{-\infty}^{+\infty} dr \ e^{-\frac{1}{2} \beta a r^2} = \frac{1 \cdot 3 \cdot 5 \cdots (2n - 1)}{(\beta a)^n} \sqrt{\frac{2\pi}{\beta a}}
\]
and the observation that all terms in the integral with an odd power of $r$ will give zero to evaluate the expanded version of the integral. The integral in the denominator of Eq. (5.3) is of particular interest because it is proportional to the local thermodynamic partition function, and so it will be shown with a few intermediate steps. We start by evaluating
\[
\mathcal{Z} = \int_{-\infty}^{+\infty} dr \ e^{-\beta \psi(r)},
\]
which will enter in all the calculations below. Expanding $e^{-\beta \psi(r)}$ as in Eq. (5.4), we have
\[
\mathcal{Z} = \int_{-\infty}^{+\infty} dr \ e^{-\frac{1}{2} \beta a r^2} \left[ 1 - \beta cr^3 + \frac{1}{2} \beta^2 b^2 r^6 + \frac{1}{2} \beta^2 b^2 r^8 - \frac{1}{2} \beta^3 b^2 cr^{10} + \frac{1}{4} \beta^4 b^4 r^{12} - \frac{1}{6} \beta^3 c^3 r^{12} + \frac{1}{4} \beta^4 b^2 c^2 r^{14} + \cdots \right].
\]
Using the results of Eq. (5.5) for each of the terms, and keeping only the lowest two powers of $1/\beta$ in the expansion gives
\[
\mathcal{Z} \approx \sqrt{\frac{2\pi}{\beta a}} \left[ 1 + \frac{1}{\beta a^2} \left( \frac{15b^2}{2a} - 3c \right) + \frac{105}{2\beta^2 a^4} (c^2 - \frac{9b^2 c}{a} + \frac{33b^4}{4a^2}) \right].
\]
For evaluating the moments $\langle r^n \rangle$, we’ll need to divide each further integral by $\mathcal{Z}$. For that purpose, we can use the approximation that
\[
\mathcal{Z}^{-1} \approx \sqrt{\frac{a\beta}{2\pi}} \left[ 1 + \frac{1}{\beta a^2} \left( 3c - \frac{15b^2}{2a} \right) \right].
\]
In this way, the lowest orders in $k_B T$ for the first four moments can be found to be:
\[
\begin{align*}
\langle r \rangle & \approx -\frac{3b}{\beta a^2} \left[ 1 + \frac{1}{\beta a^2} \left( \frac{45b^2}{a} - 32c \right) \right], \\
\langle r^2 \rangle & \approx \frac{1}{\beta a} \left[ 1 + \frac{1}{\beta a^2} \left( \frac{45b^2}{a} - 12c \right) \right], \\
\langle r^3 \rangle & \approx -\frac{15b}{\beta^2 a^3} \left[ 1 + \frac{1}{\beta a^2} \left( \frac{108b^2}{a} - 60c \right) \right], \\
\langle r^4 \rangle & \approx \frac{3}{\beta^2 a^2} \left[ 1 + \frac{1}{\beta a^2} \left( \frac{150b^2}{a} - 32c \right) \right].
\end{align*}
\]
We can then use the results of Eq. (2.50) to write the lowest four cumulants in terms of the three anharmonic force constants of the one-dimensional potential as

\[
\Delta R = C_1 \approx \frac{-3bk_BT}{a^2} \left[ 1 + \frac{k_BT}{a^2} \left( \frac{45b^2}{a} - 32c \right) \right],
\]

\[
\sigma^2 = C_2 \approx \frac{k_BT}{a} \left[ 1 + \frac{k_BT}{a^2} \left( \frac{36b^2}{a} - 12c \right) \right],
\]

\[
C_3 \approx \frac{-6b(k_BT)^2}{a^3} \left[ 1 + \frac{k_BT}{a^2} \left( \frac{144b^2}{a} - 84c \right) \right],
\]

\[
C_4 \approx \frac{(k_BT)^3}{a^4} \left[ \frac{108b^2}{a} - 24c \right].
\]

With this model for the anharmonic potential, the four cumulants can be written in terms of the three lowest order anharmonic force constants and the temperature. By using these expressions for the cumulants in the form for $\chi(k)$ given in Eq. (4.14), the effect on the XAFS of $g(r, T)$ can be modeled. In this way, an entire set of XAFS data for a sample at different temperatures can be used to measure the the one-dimensional force constants $a$, $b$, and $c$. By constraining the cumulants for each temperature to be as given in Eq. (5.11), several XAFS parameters (4 for each temperature) can be written in terms of these three “generalized variables”, significantly reducing the number of measured parameters. This not only increases the ratio of independent points to measured variables (which improves the confidence level of the measurement), but also provides a consistent set of constraints on the temperature dependence of the XAFS, thereby allowing the measurement of physical parameters that are easier to interpret and visualize than the normal XAFS cumulants.

It should be mentioned that the simple relationship between the first three cumulants pointed out by others [39, 101], namely that $\sigma^2 \approx C_3/2\Delta R$, is seen to be the case for this one-dimensional anharmonic potential.

5.2 The Thermodynamics for a Slightly Anharmonic Potential

The vibrational contribution to the classical partition function is

\[
Z = \left( \frac{m_B k_BT}{2\pi \hbar^2} \right)^{1/2} \int_{-\infty}^{+\infty} dr \, e^{-\psi(r)/k_BT}. \tag{5.12}
\]

This form for $Z$ can be easily evaluated in terms of the anharmonic force constants $a$, $b$, and $c$ for the anharmonic potential given in Eq. (5.1). The integral above is, in fact, exactly the term $Z$ developed in the previous section to evaluate the cumulants for the anharmonic potential, and we can use the results of Eq. (5.8) to simply write...
down the vibrational portion of the partition function for a two particles bound by
the anharmonic potential as
\[
Z \approx \sqrt[4]{\frac{m_R}{a \hbar^2}} k_B T \left[ 1 + \frac{k_B T}{a^2} \left( \frac{15b^2}{2a} - 3c \right) \right].
\] (5.13)

Thermodynamic properties can be derived from this partition function in the
usual ways [2, 102, 103]. The entropy, for instance, can be found using the identity
\[ S = (\partial / \partial T)_V (k_B T \ln Z), \]
which leads to
\[
S(T) \approx k_B \left[ 1 + \ln \left( \sqrt[4]{\frac{m_R}{a \hbar^2}} k_B T \right) + \frac{3k_B T}{a^2} (5b^2 - 2ca) \right],
\] (5.14)
for the entropy as a function of temperature and the force constants. Similarly, the
vibrational specific heat in one-dimension is given by
\[
C_v(T) \approx k_B \left[ 1 + \frac{3k_B T}{a^3} (5b^2 - 2ca) \right].
\] (5.15)
The three-dimensional specific heat is three times this value.

To first order, the linear thermal expansion coefficient can be approximated as the
linear expansion coefficient in the radial distance, which is measured by XAFS from
the temperature dependence of the first cumulant, \( \Delta R \) in Eq. (5.11). This gives the
linear expansion coefficient as
\[
\alpha = \frac{L(T) - L_0}{TL_0} \approx \frac{\Delta R}{TR_0} \approx \frac{-3bk_B}{R_0 a^2},
\] (5.16)
where \( L \) is a linear dimension. This simple estimate will almost certainly underesti-
mate the linear expansion coefficient\(^2\). This simple estimate of the linear expansion
coefficient (which is really the radial expansion) can be improved by evaluating the
expansion coefficient of volume, \( \alpha_V = (V(T) - V_0)/(TV_0) \), which is simply 3 times
the linear expansion coefficient for cubic materials. The volume expansion coefficient
can be easily described by just the change in radial component, by averaging
\( R^3 = (R_0 + r)^3 \). Thus,
\[
\alpha_V = \frac{V(T) - V_0}{TV_0} = \frac{\langle R(T)^3 \rangle - R_0^3}{TR_0^3},
\] (5.17)
\(^2\) The radial component is always less than or equal to any of the three axial components. Since
the observed change in radial distance will ignore the movements of the atoms perpendicular to
the radius, this will always underestimate the axial expansion.
where \( R = R_0 + r \), and the average \( \langle R^3 \rangle \) is evaluated as the usual expectation value,

\[
\langle R^3 \rangle = \frac{\int dr (R_0 + r)^3 e^{-\psi(r)/k_B T}}{\int dr e^{-\psi(r)/k_B T}}.
\] (5.18)

Using this expression and the same basic formalisms used in section 5.1 to determine thermodynamic averages, the linear expansion coefficient can be easily evaluated in terms of \( R_0 \) and the three anharmonic force constants \( a, b, \) and \( c \). The result is

\[
\alpha = \frac{\alpha V}{3} \approx \frac{k_B (a - 3R_0b)}{R_0^2a^2}.
\] (5.19)

The difference between this more elaborate evaluation of the linear expansion coefficient and the simpler version in Eq. (5.16) can be significant, as will be shown in chapter 6.

It should be pointed out that, according to Eq. (5.19), there \textit{will} be linear expansion even if the potential \( \psi(r) \) is harmonic (so that \( b = 0 \)). This is a consequence of \( g(r, T) \propto \exp(-\psi(r)/k_B T) \) giving the probability of an atom being a distance \( R = R_0 + r \) from another atom, and is in contrast to the more usual description of a potential for an atom around a particular lattice site, which will only give lattice expansion if there is an anharmonic contribution to the potential. For a harmonic \( \psi(r) \), the atom samples more distances as the temperature increases in such a way that the average distance (\( = R_0 \)) remains constant. But because the sampled distances do not contribute equally to the average volume, the average volume of the interatomic bond will change as the atom samples more distances at higher temperatures. In fact, longer distances tend to increase the volume more than shorter distance tend to decrease the volume, so that the volume will increase even for a harmonic \textit{radial} oscillator.

The Grüneisen parameter, a dimensionless parameter that relates the anharmonicity in a crystal to the pressure-dependence of the elastic constants and the vibrational energy, is usually defined as \( \gamma = d \ln \nu / d \ln V \), where \( \nu \) is the principle vibrational frequency of the solid and \( V \) is the volume. In one-dimension, this can be approximated as \( \gamma = d \ln [\psi''(r)] / 6d \ln r \). Writing the three-dimensional Grüneisen parameter as three times this (which should be a reasonable approximation for cubic systems where all three degrees of freedom are equivalent), and evaluating this at the equilibrium position \( R_0 \), gives

\[
\gamma \approx \frac{3(a + 2bR_0)}{2a + 3bR_0}.
\] (5.20)

As a final thermodynamic parameter, we discuss the isothermal bulk modulus, which is related to many of the above parameters. Thermodynamically defined as \( K = \)
\(-V(dP/dV)_T\), the bulk modulus can be related simply to the heat capacity and the Grüneisen parameter as \(K = C_\gamma / V\alpha\). Since we’re evaluating these thermodynamic properties based on the interatomic bonds, the volume in this expression is the volume \textit{per bond}, which is 1/6 the volume per atom (in an fcc material, each atom is bonded to twelve neighbors, so that the volume per atom contains twelve half-bonds). We could calculate \(K\) for a three-dimensional solid, so that the \(C_\gamma\) used is three times that given in Eq. (5.15), and the \(\alpha\) used is \(\alpha V\), but it makes no difference to the final result, which is

\[
K \approx \frac{2\pi a^2(a + 2bR_0)}{2\pi R_0(a - 3R_0b)(2a + 3bR_0)}.
\]

Before leaving this discussion to pursue applications to actual data, and for completeness, we present the standard Einstein model for the mean-square disorder \(\sigma^2\) in an XAFS experiment. This is discussed in greater detail in many sources [30, 36, 34], and the basic connection with thermodynamics is that \(\sigma^2\) is related to the vibrational frequency. The Einstein model gives

\[
\sigma^2 = \frac{\hbar^2}{2M_R k_B \theta_E} \coth \left( \frac{\theta_E}{2T} \right),
\]

where \(M_R\) is the reduced mass of the two bonded atoms, and \(\theta_E\) is the Einstein temperature. In the classical limit (where this work dwells), where \(T > \theta_E\), \(\coth(\theta_E/2T)\) can be approximated as \(2T/\theta_E\), so that

\[
\sigma^2 \approx \frac{\hbar^2 T}{M_R k_B \theta_E^2} \quad T > \theta_E.
\]

Comparing this to the expression for \(\sigma^2\) in Eq. (5.11), we see that the Einstein temperature can be evaluated in terms of the force constant \(a\) as

\[
\theta_E \approx \frac{\hbar}{k_B} \sqrt{\frac{a}{M_R}}.
\]

The basic scheme to use XAFS to get thermodynamic measurements is to write the cumulants for the XAFS equation in terms of the anharmonic force constants \(a\), \(b\), and \(c\) according to Eq. (5.11). These force constants can then be determined by finding which values give the best-fit to temperature-dependent XAFS data. The \textit{local} thermodynamic properties developed in this section can then be evaluated from the XAFS-measured force constants, and compared to values by other (more traditional thermodynamic) techniques.