Data Analysis and Energy Calibration

In the last chapter we discussed how the NaI and Ge detectors work. Here we discuss how one can analyze the data from these detectors. We are mainly interested in the photopeaks in our NaI and Ge spectra. More specifically, we want to determine the channel number of the center of the peak and the area under the photopeak. The channel number of the peak center will tell us the energy of the gamma. The area under the photopeak gives information about the number of gammas emitted (with the particular energy of the peak center). Thus, from our gamma spectra we will be able to identify the isotopes in a sample, and determine how much of the isotope is present.

We start by discussing peak fitting, which will be used to accurately measure the center of the peak and the area under the peak. We will then describe methods used for energy calibration and determing the efficiency of the detector. Although the analyses presented here apply to gamma spectra, the same techniques are used in nearly all spectral analysis.

Gaussian Peak Fitting

Curve fitting is a standard technique in which the parameters of a fitting function are varied to best describe the data. The photopeaks for the NaI and Ge detectors are well described by a Gaussian, or Normal, distribution. This can be seen in the figure shown.

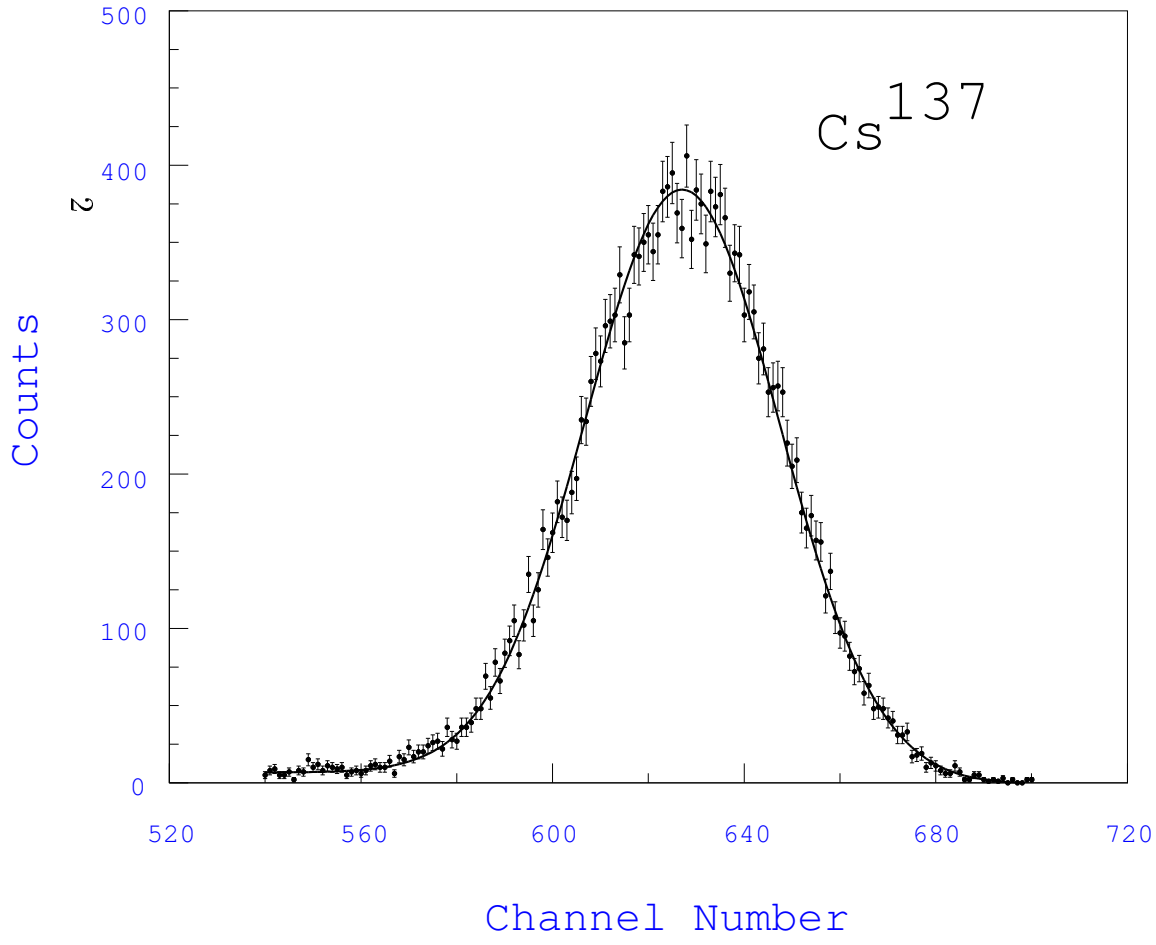
Gaussian fit to photopeak

The solid curve is a Gaussian plus background fit to data taken from the NaI MCA/detector system that is used in laboratories. The error bars are equal to the square-root of the counts in the various channels.

The approach taken in curve fitting is to assume that the data follow a certain function which contain a number of unknown parameters. Then the parameters are varied to "best fit" the data. We will be fitting the photopeaks in our gamma spectra, and assume that the photopeak has the shape of a Gaussian function due to the gamma ray plus a background. We will limit our fit to the data around the photopeak where these assumptions apply. A Gaussian function plus background can be described by 5 parameters (h, C_0, σ, b_1, and b_2):

$$y(C) = he^{-((C-C_0)/\sigma)^2} + background(b_1, b_2) \tag{1}$$

where C in the channel number, C_0 is the peak center, h is the height of the Gaussian function, and σ is related to the width of the Gaussian shape. The function $Y(C)$ represents the number of counts in channel C for the theoretical fitting function.


Cs137
Counts
Channel Number
0
100
200
300
400
500
520
560
600
640
680
720

The background function that we use is a flat plateau before the peak of height b_1, a flat plateau after the peak of height b_2, and a line connecting the two plateaus. The plateau before the peak stops at channel $C_0 - 2\sigma$, and the plateau after the peak starts at channel $C_0 + 2\sigma$. So the line starts at channel number $C_0 - 2\sigma$ with height b_1, and ends at channel number $C_0 + 2\sigma$ with height b_2.

How the Best Fit is determined

The "best fit" to the data is determined by varying the 5 parameters in the fitting function Y(C) so that Y(C) comes as close to the data as possible. Mathematically this is accomplished by defining a chi-square function, χ^2, as follows:

$$\chi^2 = \sum_{C=C_i}^{C=C_f} \left(\frac{y(C) - Exp(C)}{\sqrt{Exp(C)}}\right)^2 \tag{2}$$

where $Exp(C)$ is the experimental value for the number of counts in channel C. The statistical uncertainty of $Exp(C)$ is $\sqrt{Exp(C)}$ from our analysis of statistical uncertainty. Thus the uncertainty squared is just $Exp(C)$, which is the denominator in the fraction above. For a particular channel C, $(y(C) - Exp(C))$ is just the difference between the fitting function and the data. One squares this difference, to make it positive, then divides by the uncertainty squared. The chi-square function is just the sum of the sum of the squares of the difference between the fitting function, $y(C)$, and the data divided by the error from an initial channel C_i to a final channel C_f. The smaller the value of the χ^2 function, the better the curve $Y(C)$ fits the data.

The function $y(C)$ and hence χ^2 contain 5 parameters: h, C_0, σ, b_1, and b_2. The "best fit" is determined by finding values for these 5 parameters which make χ^2 as small as possible. When the function χ^2 is minimized, the curve $y(C)$ will be as close to the data as possible. This technique is called chi-square minimization, and is used in many areas of data analysis.

In the laboratory, a computer program will do all the calculations for us. We will only need to set a curser for the initial channel C_i and final channel C_f for the Gaussian fit. We will refer to the range of channels between C_i and $_f$ as the window for the fit. The instructor will describe how the program works. The user just presses the appropriate keys, and the program varies the 5 parameters using a grid search to find values that make the χ^2 function as small as possible. The user just keeps pressing the automatic search key until the χ^2 function stops decreasing.

How small should χ^2 be? It is best to divide χ^2 by the number of data points. This number is referred to as the chi-square per data point, and tells us how many

standard deviations (on the average) the fit is away from each data point. The chi-square per data point, χ^2/N, should be less than 2.0 for an acceptable fit. Ideally χ^2/N should be between 1.0 and 1.5. The computer program will print the χ^2/N to let the user know the quality of the fit. The users main task is to supply the channel window for the fit, C_i and C_f. You want to be sure that you include enough of the photopeak and flat background, but not too much extraneous background in choosing the window for the fit. Your final results should not be too sensitive (hopefully) to the choice of channel window.

Two parameters from the fit will be of interest to us: the peak center C_0 and the area, A, under the peak. The accuracy which we can extract these from the data depend on our knowledge about the shape of the peak and background. Fortunately the peaks are very close to a Gaussian function, and the background is well parameterized by the model above. Next we discuss how to calculate the resolution of the detector, and how to calibrate the energy scale.

Detector Resolution

The resolution of a detector is a measure of how narrow the peaks are. A commonly used measure is the "**F**ull **W**idth at **H**alf **M**aximum", FWHM. It is just the width of the photopeak at which the values are 1/2 the maximum value. Let C_0 be the channel number of the center of the photopeak, where the maximum height is h. Let C_+ be the channel number greater than C_0 where the counts are $h/2$. Let C_- be the channel number less than C_0 where the counts are $h/2$. Thus, $C_+ - C_-$ is the width of the photopeak where the counts are 1/2 the maximum value. The FWHM is defined as follows:

$$FWHM = \frac{C_+ - C_-}{C_0} x100\% \tag{3}$$

One can scan through the channels and use the above formula for the FWHM. However, a better way is to use our results from a Gaussian χ^2 fit to the peak. Channel number C_+ satisfies the equation:

$$\frac{h}{2} = he^{-(C_+ - C_0)^2/\sigma^2} \tag{4}$$

Solving this equation for C_+ gives:

$$C_+ = C_0 + \sigma\sqrt{ln2} \tag{5}$$

Similarily, one obtains for C_-

$$C_{-} = C_0 - \sigma\sqrt{ln2} \tag{6}$$

Using the formula for the FWHM, we have

$$FWHM = 2\sqrt{ln2}\,\frac{\sigma}{C_0} \tag{7}$$

Using Gaussian curve fitting one can obtain an accurate value for the FWHM by using this formula.

Energy Calibration of Solid Scintillation Detectors

As discussed in the last section, the channel number of the photopeak is approximately proportional to the energy of the gamma (or X-ray) particle. The scaling factor is controlled by the amplifier gain. If we assume an exact linear relationship between channel number and energy, we have:

$$E = aC \tag{8}$$

where E is the energy of the gamma and C is the channel number of the center of the photopeak. The constant a is a scaling factor with units of energy/(channel number). To determine the scaling factor a, we need a source that produces a gamma particle with a known energy. Then by measuring the channel number of this gamma, we can determine the constant a. Once a is known, the energy of an "unknown" gamma can be determined from the channel number of the photopeak, C, and the equation above.

In practice, one usually uses more than one standard source for the energy calibration of the detector. The channel numbers of three or four gamma particles of known energy are measured. Then a "best fit" line, through these "standard" data points, is used as a calibration line. The calibration line might not pass through the origin, since the detector system might have an offset.

Using Gaussian curve fitting, one is able to measure the channel number of the center of the photopeak, C_0, very accurately. Using our 1028 channel detector system, NaI, C_0 can be measured to within 0.2 channel numbers. With this accuracy, deviations from the simple linear relation between E and C can be observed. One could chose a quadratic fitting function: $E = aC + bC^2 + c$ or a higher order polynomial fit. However, the gain of the amplifiers used in the laboratory can vary in time. After turning on the amplifier and power supply, one should wait 10 minutes for so for the electronics to "warm up". The drift in amplifier gain is often related to the quality of the amplifier/power supply system. With our less expensive amplifier

systems, the drift in amplifier gain is important for accurate energy determination. Using Gaussian curve fitting, one can observe the center of the photopeaks drift in time. The drift can be as high as one channel number in 30 minutes. With our more expensive amplifiers, the drift is minimal and does not play a big role.

It is our experience that the best approach (for the most accuracy) is to use a linear fitting form and to adjust for the drift in the amplifier gain. To compensate for amplifier drift, one can use the calibrated sources before and after a measurement. By taking an appropriate average of the center of the calibration photopeaks, one can improve the accuracy of the energy determination significantly.

When running a series of experiments, usually the amplifier gain is set and not changed. Before an experiment, one needs to determine the highest gamma energy that you will measure and set the gain so that this high energy gamma produces a photopeak at a channel number near the right of the screen. For example, in our experiments we usually measure gammas with energies up to 1500 KeV. If we are using a 1028 channel system, we want the highest gammas to be at channel number say 950. This results in a calibration constant of roughly 1500/950 = 1.6 KeV/(channel number). For the best accuracy in a series of experiments, one would like to have the photopeaks of the isotopes under investigation span the whole range of channel numbers.

Efficiency Calibration of Solid Scintillation Detectors

The efficiency ε of a detector is defined as (the number of particles detected)/(the number of particles emitted):

$$\varepsilon = \frac{the\ number\ of\ particles\ detected}{the\ number\ of\ particles\ emitted} \tag{9}$$

The efficiency is a number between zero and one. If we know the efficiency of our detector, then measuring the number of particles detected will allow us to determine the number of particles emitted in our sample. The efficiency of a detector will depend on a few factors, the most important are:

1.**The source-detector geometry**: The number of particles detected will depend on how close the source is to the detector. The closer the source is to the detector, the larger the efficiency will be.

2.**The size of the detector**: Larger detectors will usually be more efficient, since they have a larger volume for the particles to absorbed in.

3.**The energy of the gamma (or X-ray) radiation**: The photopeak is produced by photo-absorption. The photo-absorption process has a strong energy dependence. For high energy photons, photo-absorption has a lower probability to occur than photons of low energy.

For solid scintillation detectors, NaI and Ge, the dependence of ε on energy, number 3 above, is quite large. For example, NaI detectors can detect 100 KeV gammas about 4-5 times more efficiently than 1200 KeV gammas. This means that although a photopeak at 1200 KeV is small compared to one at 100 KeV in a particular spectrum, there might be more 1200 KeV gamma emitted than 100 KeV gammas.

Since the efficiency depends on the three factors listed above, one often keeps the source-detector geometry fixed during a series of experiments. That is, for a series of experiments one places all the samples in the exact location relative to the detector. Also, one uses samples that are all the same size and shape. If this is done, then factors 1 and 2 above are the same for all the samples in a particular experiment. In this case, the only efficiency calibration necessary is the energy dependence of ε. The energy dependence for a particular source-detector geometry is measured by using standardized sources. One can purchase sources in which the activity has been calibrated by the manufacture. If the activity of the source is known, then the number of gamma particles emitted can be calculated. By measuring the number of gammas (of a particular energy) detected during a specific time interval, the efficiency ε can be determined.

An Example without including a Geometry Correction

As a sample efficiency calculation neglecting source-detector geometry, we consider the 662 KeV gamma emitted by Cs^{137}. Suppose the Cs^{137} source was 1.1μCi on May 20, 1990. This value is written on a calibrated sample. We place this source under our detector and record counts for 1 minute. Suppose the area under the 662 KeV photopeak is 20,000 counts. First one calculates the activity today, which is (say) May 20, 2004:

$$A = 1.1\mu Ci \frac{37000 decays/sec}{\mu Ci} (\frac{1}{2})^{(14 years)/(30 years)} \tag{10}$$

$$A = 30845 \frac{decays}{sec} \tag{11}$$

since the half-life for Cs^{137} is 30 years, and 1μCi equals 37,000 decays/sec. Next one calculates the number of gammas that are emitted per second.

$$\gamma's\, emitted = 30845(0.85(\frac{\gamma's}{decay})) = 26200\frac{\gamma's}{sec} \tag{12}$$

since the gamma yield factor is 0.85. In one minute, (26218 gammas/sec)(60 sec/min) $= 1.57x10^6$ gammas are emitted. The efficiency for the detector to detect a 662 KeV gamma is then

$$\varepsilon = \frac{20000 counts\, recorded}{1.57x10^6 \gamma's\, emitted} = 0.0127 \tag{13}$$

The number 0.0127 is the efficiency of the detector for 662 KeV gamma particles for *the particular source-detector geometry* that we have used. Note that efficiency is unitless.

Efficiency Calculations including Source-Detector Geometry

The distance from the sample to the detector and the size of the detector are important factors that affect the detector's efficiency. We would like to include these affects in our measurements. The method that seems to work best (i.e. is simple and fairly accurate) with our NaI detectors is to use the ansatz:

$$\frac{Counts\, Detected}{time} = \frac{\gamma's\, emitted}{time}(\frac{A}{4\pi(x+d)^2})\,\epsilon \tag{14}$$

where A is the cross sectional area of the detector, x is the distance from the source to the bottom of the detector, and d is the distance from the bottom of the detector to the "effective center" of the detector. The geometry factor $A/(4\pi(x+d)^2)$ represents the fraction of gammas emitted that go through the detector. This geometry factor is just an approximation, but usually gives consistent results in our experiments. In our laboratory, we have NaI detectors of two sizes: with a diameter of 1 1/2 inches, and with a diameter of 2 inches. One can measure x, so once the "effective distance" d is known, the efficiency ϵ can be determined.

In our laboratory session, you will determine d by collecting data from one source located at different distances x from the detector. A computer program that does a χ^2 fit of the data to determine the best estimate of d is available on the lab computers. The program was written by Sue Hoppe (2003), a Cal Poly Pomona physics major.

Once d is known, we will measure the efficiency ϵ for different gamma energies. For standards, we will use Cs^{137} ($E_\gamma = 662$KeV), Mn^{54} ($E_\gamma = 835$KeV), and Na^{22} ($E_\gamma = 511$ and 1275 KeV). When a graph of efficiency vs. energy, $\epsilon(E)$ is plotted, the strong energy dependence of ϵ can be demonstrated. The energy dependence of

ϵ is related to the energy dependence of the photo-absorption process. For gamma energies between 100 KeV and 1500 KeV, the efficiency tends to follow a power law relationship with energy.

Self-Absorption Corrections:

If the sample being measured is large, then self-absorption needs to be considered. Self-absorption means that the sample itself absorbs some of the radiation. The problem with self-absorption is that the sample absorbs gamma particles with different energies differently. Low energy gammas are absorbed by the sample more readily than high energy ones. For large samples, one has to do an efficiency calibration with standards that are as large as the sample. In our laboratory, the only large samples that we measure are environmental samples, e.g. soil and other bulk samples. For these samples we have a nice way to calibrate the efficiency of the detector. The method is to graph the counts recorded divided by the yield, $Counts/Yield$ as a function of gamma energy. A log-log graph of these quantities plots in a straight line over the energy range 100KeV to 1500KeV. These properties are well suited for environmental samples which contain the U^{238} and/or the Th^{232} decay series. We will discuss these topics in our experiment on natural radiation.

Energy and Efficiency Calibration of the Liquid Scintillation Detector

As with the gamma detectors, the channel number recorded by the detector is approximately proportional to the energy of the radiation. The energy of the beta particle given off in beta decay is shared with the neutrino. Hence the beta spectrum does not have a nice photopeak as with gamma decay. However, the maximum energy that the emitted beta can have is known and can be looked up in the literature. For ^3H, E_{max} = 18.6KeV, and for ^{14}C, E_{max} = 156KeV. By measuring the maximum channel number for which there are counts, one can calibrate the energy per channel number.

For the liquid scintillation detector, the sample is in the detecting material. Hence, the efficiency of a liquid scintillation detector is high, and can be between 50% and 100%.

Quenching and Corrections

Although the liquid scintillation detector "detects" nearly every beta emitted, all the energy of the beta is often not recorded. As the beta particle is slowed down by the fluor, not all of its energy goes into exciting the fluor. This is known as quenching.

Quenching can occur if the solvent absorbs the energy of the beta but does not give off light, or the solvent absorbs the light.

The amount of quenching may depend on how the sample is prepared, e.g. how much water is added, etc. Quenching will affect the shape of the spectrum, and can affect the efficiency. **When measuring many biological samples, you need to be sure that each sample is being counted with the same efficiency and consequently the same amount of quenching**.

A common method to insure that all samples are quenched the same is called the "channels-ratio method". In this method, two windows are chosen. One then divides the number of counts in one window by the number of counts in the other window. If the two windows are chosen correctly, then the ratio of the counts in the two windows depends strongly on the shape of the spectrum. If each sample is quenched the same, then the ratio of the counts in the two windows is the same for the samples. By examining the ratio of (the counts in one window)/(the counts in the other window) one can check if the quenching is different in any of the samples.

For the channels-ratio method, one should choose the windows such that half of the total counts are in one window, with the other window containing the remainder of the counts. This way the ratio should be close to one, and any quenching should be easy to detect.