Data Analysis and Energy Calibration

In the last section we discussed how the NaI and Ge detectors work. Here we focus on analyzing 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. The efficiency calibration of the detector will be covered in the next experiment.

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. The quality of the fit is demonstrated in the figure.

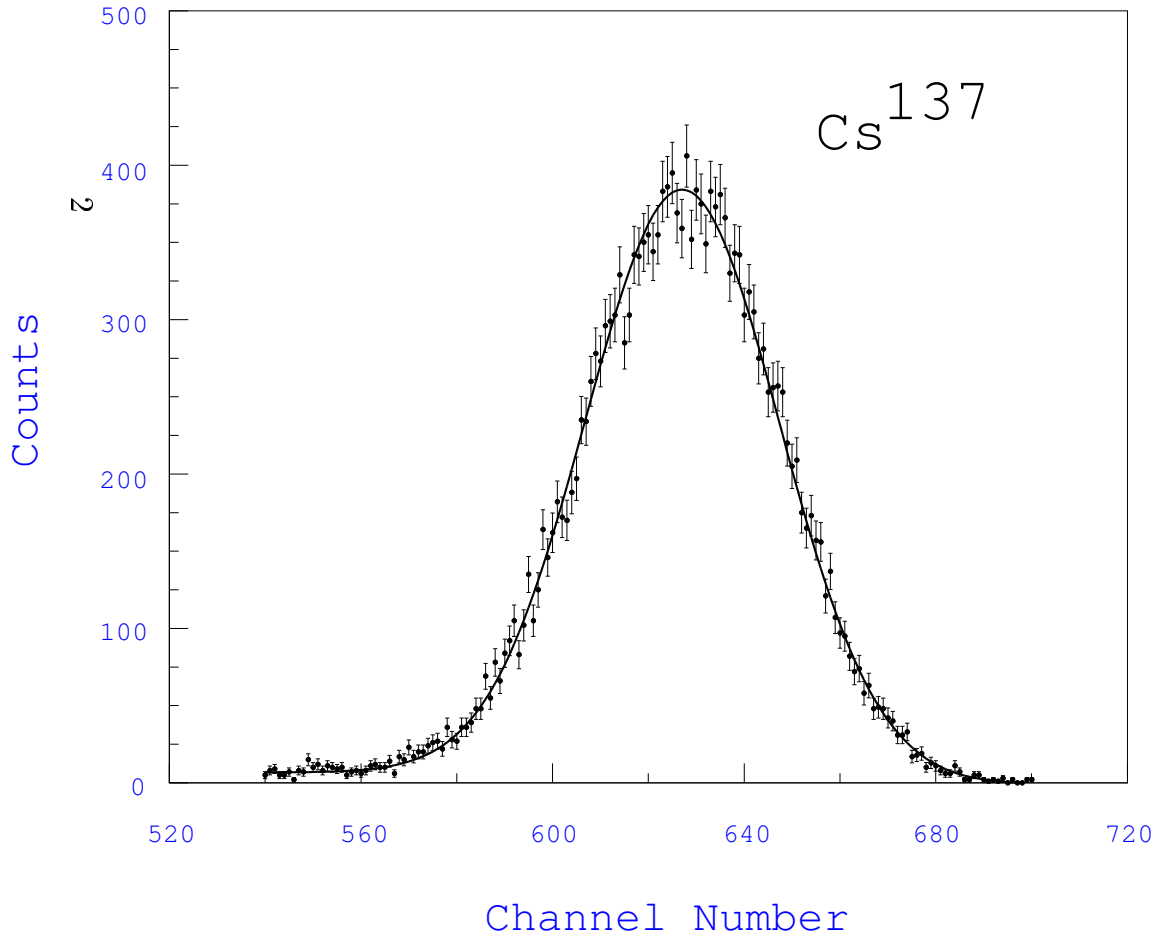
The solid curve is a Gaussian plus background fit to data taken from the NaI MCA/detector system that is used in laboratories. The error bars are equal to the square-root of the counts in the various channels.

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.

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.


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

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.