Aspects of Data Treatment for Transverse µSR

B.D.Rainford

Department of Physics & Astronomy, University of Southampton,
Southampton SO17 1BJ, England

Introduction

My main message is that in µSR experiments we are always fighting with poor counting statistics, so we must do nothing to our data which makes matters worse, and we should do whatever we can to extract the most from what we have. The approach is geared to the typical set-up encountered at the pulsed muon beams at ISIS, but is readily generalised to other situations.

1. A typical ISIS µSR data set.

A data set collected on the MUSR or EMU spectrometers at ISIS consists typically of 32 histograms containing between 1000 to 2000 time channels; the channel widths are 8 or 16 ns. These widths are chosen to match the width of the muon pulse, which is of order 70ns (FWHM). As an example let us consider a data set of 1500 channels, with channel width 16ns. The total time window for data collection is then 0 to 24 µs. As the muon lifetime τ_μ is 2.2 µs, the end of the time window corresponds to $10.9\tau_\mu$, by which time the initial decay count rate will have dropped by a factor exp (-10.9) = $2\ .10^{-5}$! Fortunately the background at ISIS is very low, so it is still possible to collect data out to these times, though the statistics will inevitably be poor.

To develop a feeling for counting statistics, let us consider a standard silver run, measured for one hour (this is a typical duration for a measurement), corresponding to 2.10^7 events (20"Mev"), summed over all 32 histograms. The counts in the n'th channel of the j'th detector are given by:

$$C_j\,(t_n) = N_j^0\,\Delta t\,\exp(-t_n/\tau_\mu)\,[1 + A_j \cos(\omega_L t_n - \phi_j)] \qquad (1)$$

where A_j is the asymmetry parameter (typically A_j = 0.23 at ISIS), ϕ_j is the phase, Δt is the channel width, N_j^0 is the initial count rate, $\omega_L = \gamma B$ is the muon Larmor precession frequency for an applied field B and the gyromagnetic ratio $\gamma/2\pi$ = 135.54 MHz/tesla. Ignoring the oscillatory term in (1) for the moment, we can sum over all time channels, to estimate the total number of events per histogram:

$$\sum_n C_j(t_n) \approx \int_0^\infty N_j^0 \exp(-t/\tau_\mu)dt = C_j(0)\tau_\mu \,/\, \Delta t$$

Taking the total counts to be $2.10^7/32$, then we find that $C_j(0)$, the counts at the start of each histogram, to be about 4500. The statistical error in this count is $\sqrt{C_j(0)}$ = 67. Of all the events, we are only interested in that portion associated with the asymmetry, $A_jC_j(0)$ = 0.2 x 4500 = 900 counts. It follows that the statistical error in the asymmetry, at the *beginning* of the histogram, where we have the highest count-rate, is only 67/900 ≡ 7.4%. This is not very impressive! And at longer times it will be worse. If we wanted, say, 1% statistics on the initial asymmetry we would need $\sqrt{C}/(0.2C) = 0.01$, corresponding to $C_j(0) = 2.5.10^5$ counts, or a total of 10^9 events in all 32 histograms! This would mean running for 55 hours on ISIS!

The moral is that in the average μSR experiment we are dealing with quite poor counting statistics. This is largely a consequence of (i) the signal being associated with only 20% of the total counts, and (ii) the short lifetime of the muon. Clearly it is essential, when analysing μSR data, to make every event count and to avoid any process which throws away information, e.g. grouping detectors with different phases together; this might be useful for summary purposes, or plotting data during an experiment, but it will reduce the effective asymmetry, and is therefore *a bad thing* for the ultimate analysis of μSR data.

2. Determination of instrumental parameters.

2.1 Phases and start time.

The time t_n in (1) is measured with respect to the start time t_0, when the muons arrive in the sample: $t_n = n\Delta t + t_0$. The effective phase of the oscillatory term is therefore $\psi_j = \phi_j - \omega_L t_0 = \phi_j - \gamma B t_0$. All detectors should have a common value of t_0. This can be determined by measuring silver spectra for several different transverse fields B, fitting the count-rate to (1) to determine the effective phases ψ_j, then plotting these as a function of B. The plots give straight lines which, when extrapolated to zero field, give the true phases ϕ_j. The start time t_0 can be extracted from the gradients. Alternatively the phases ψ_j can be determined for different assumed values of t_0, then plots of $\psi_j(t_0)$ for different fields should intersect at the true value of t_0.

Fitting of (1) to a silver spectrum in order to determine the phases is straightforward: it can be implemented by a linear least squares method, since (1) can be written as: $C_j^t = \alpha_j F_t + \beta_j G_t$, with $\alpha_j = \sin \phi_j$ and $\beta_j = \cos \phi_j$. Then following the usual approach

$$\chi^2 = \sum_t [C_j^t - \alpha_j F_t - \beta_j G_t]^2 / \sigma_{jt}^2$$

and the values of α_j and β_j can be found by minimising χ^2 in the normal way.

2.2 Dead time corrections

The count-rate equation (1) assumes a linear detector response, i.e. that the number of detected events in a given time channel is simply proportional to the number of decay positrons at that time. In general this is a reasonable assumption for low count rates, but at high count-rates we have to correct for the "dead time" of the detector. This arises because the counting chain cannot respond to two events that come very close together: the counter is "dead" for a time τ_d after the first positron is detected. Since for μSR data the count rate is initially high and then falls with time, the effect of the dead time will be biggest at early times, and will become negligible at later times. The corrections for dead time can be derived from a standard zero field silver run, preferably one with good statistics.

To see the effect of dead time, let us consider a simple case, where N(t) represents the ideal counting rate and M(t) is the recorded counting rate. In a single time channel of width Δt the detector will be dead for a time M(t) $\Delta t\ \tau_d$, so the measured count in this time channel will be M(t) Δt = N(t) Δt [1 - M(t) τ_d]. It follows that the recorded count rate is

$$M(t) = N(t) / [1 + N(t)\ \tau_d] \quad (2)$$

Clearly when N(t) is small, and the product N(t)τ_d «1, then M(t) ≈ N(t). However when the count-rate is very high, the measured counts tend to the limiting value M(t) = 1/τ_d. To determine the dead times for each detector, the reciprocal of the measured counts in a zero field silver run can be plotted against the reciprocal of N(t) with N(t) = N_0 exp (-t /τ_μ). From (2) it can be seen that a straight line should result with intercept τ_d:

$$1/M(t) = 1/N(t) + \tau_d$$

The value of the dead time can then be extracted from a least squares fit to the straight line. If only a transverse field silver run is available, the same procedure can be used, but in this case the ideal counts are as given in equation (1). In practice for a typical ISIS count rate the dead time corrections at short times are about 5% or less.

3. Analysis of μSR spectra

There are two perspectives for the interpretation of transverse field μSR spectra:

(a) the muon spins precess at a unique frequency ω_L, but their polarisation decays with time due to the muons' interaction with static or dynamic fluctuations in its environment. In metals the precession frequency ω_L might be shifted slightly from the value γB_{ext} (Knight shift), reflecting an enhanced susceptibility at the muon site in the sample.

(b) the muons at different sites see a distribution of internal fields, and therefore precess at different rates. This is the case, for example, in the mixed state of superconductors.

We will consider these situations in turn in sections 3.1 and 3.3, following. In both cases there are special techniques available which are useful for tackling data comprising multiple histograms.

3.1 Modelling the depolarisation in the time domain.

Case (a) If the muons precess at a common frequency ω_L, but their polarisation decays with time, the counts in the n'th time channel of the j'th detector can be written:

$$C_j(t_n) = N_j^0\ \Delta t\ \exp(- t_n / \tau_\mu)\ [1 + A_j\ P_x(t_n) \cos(\omega_L t_n - \phi_j)] \quad (3)$$

$P_x(t)$ describes the decay of the muon polarisation. The form of this decay depends on the physics of the muon's interaction with its environment. In the limit of fast temporal fluctuations of the local field or of rapid hopping of the muon from site to site it can be shown that $P_x(t) = \exp(-\lambda\, t)$: this is traditionally (and confusingly) called "Lorentzian damping" because this limit is identical to the limit of strong motional narrowing in NMR, where the line-shape (equivalent to the Fourier transform of $P_x(t)$), is indeed a Lorentzian function. In the limit of slow temporal fluctuations, but where there is a Gaussian distribution local fields (inhomogeneous broadening) it is readily shown that $P_x(t) = \exp(-\sigma^2 t^2)$. Other functions in the literature, e.g. the Abragam function, are designed to describe circumstances between these two limiting cases. The point is that the shape of the envelope of the decaying polarisation gives information about the physics of the interaction at the muon site, which can be described by an analytical function with a few parameters (e.g. λ or σ in the above expressions). Analysis of the data is then a simple matter of least squares fitting the expression (3) to the data, with the appropriate form of $P_x(t)$. Following our dictum in section 1, we should, if possible, fit *all* the histograms (j=1 to 32) at the same time, without grouping them together, after making necessary corrections for dead times. In this case there will be three instrumental parameters for each histogram, namely N_0^j, A_j and ϕ_j, as well as the parameters describing the polarisation $P_x(t)$. In practice it is best to predetermine the phases as described in section 2.1, using a standard silver run, since the phases might not be well determined in a fit to the data, especially when the polarisation is rapidly damped. Fitting 32x1500 time channels with 65 parameters might seem like a tall order, but computing power is cheap compared to the cost of doing μSR experiments. In the next section we show a way of combining multiple histograms without losing information, which reduces considerably the effort of fitting models to the data.

3.2 Combining multiple histograms.

We need a technique that allows us to condense a data set comprising multiple (e.g. 32) histograms without losing information. The following method is based on unpublished work by G.J.Daniell and has been implemented in a practical program for ISIS data by W.R.Dickson. Starting from the raw data, as described by (3) above, it is straightforward to extract the exponential decay term due to the muon lifetime and the uniform contribution. This yields the component of the signal that oscillates about zero:

$$d_j(t_n) = N_j^0\, \Delta t\, A_j\, P_x(t_n) \cos(\omega_L t_n - \phi_j) \;=\; D_j(t) \cos(\omega_L t_n - \phi_j) \qquad (4)$$

for each histogram, j = 1 to J and time channel, n = 1 to N. We also have the corresponding variances $\sigma_j(t_n)$ in the $d_j(t_n)$. Now we can project all the data onto a given phase by taking a linear combination of all the histograms. Suppose we choose a value of zero for the phase: we can define a function X(t) such that

$$\begin{aligned} X(t_n) &= D_j(t_n)\cos(\omega_L t_n) = \sum_j x_j d_j(t_n) \\ &= \sum_j x_j D_j(t_n)\{\cos(\omega_L t_n)\cos(\phi_j) - \sin(\omega_L t_n)\sin(\phi_j)\} \qquad (5) \end{aligned}$$

Obviously this identity is true if the weights x_j are chosen so that $\sum_j x_j \cos(\phi_j) = 1$ and

$\sum_j x_j \sin(\phi_j) = 0$. We can use these two conditions as constraints in a least-squares minimisation of the square of the variance of our new histogram $X(t_n)$ by using Lagrange multipliers. We then have to minimise

$$\{\sum_j \sum_n x_j^2 \sigma_j^2(t_n) - \lambda[\sum_j x_j \cos(\phi_j) - 1] - \mu \sum_j x_j \sin(\phi_j)\}$$

with respect to the weights x_j , where λ and μ are the Lagrange multipliers. This leads to the result:

$$x_j = [\lambda \cos(\phi_j) + \mu \sin(\phi_j)] / \{2\Sigma_n \sigma_j^2(t_n)\} \qquad (6)$$

We can now use the two constraints equations to determine λ and μ. In order to simplify the resulting expression let us introduce the following notation for the sums involved:

$$\Sigma^{ss} = \sum_j \sin^2(\phi_j)/\Delta_j \; ; \qquad \Sigma^{sc} = \sum_j \sin(\phi_j)\cos(\phi_j)/\Delta_j \; ; \qquad \Sigma^{cc} = \sum_j \Sigma_j \cos^2(\phi_j)/\Delta_j$$

where $\quad \Delta_j = 2 \sum_n \sigma_j^2(t_n)$

The final result for the weights x_j is then:

$$x_j = \Delta_j^{-1}\{ \cos(\phi_j) - \sin(\phi_j) \Sigma^{sc} / \Sigma^{ss}\} / \{ \Sigma^{cc} - (\Sigma^{sc})^2 / \Sigma^{ss}\} \qquad (7)$$

We can now choose a second orthogonal value for the phase and define another function Y(t) such that:

$$Y(t_n) = D_j(t_n) \sin(\omega_L t_n) = \sum_j y_j d_j(t_n) \qquad (8)$$

The weights y_j are now chosen so that $\sum_j y_j \cos(\phi_j) = 0$ and $\sum_j y_j \sin(\phi_j) = -1$. Following the same method we find that the weights are given by:

$$y_j = \Delta_j^{-1}\{ -\sin(\phi_j) + \cos(\phi_j) \Sigma^{sc} / \Sigma^{cc}\} / \{ \Sigma^{ss} - (\Sigma^{sc})^2 / \Sigma^{cc}\} \qquad (9)$$

By combining all J histograms with the appropriate weights x_j and y_j we can project out two histograms with orthogonal phases. These contain all the original information and the variance of each is of order 1/J times that of a single histogram. These projections of the data are very convenient for displaying the "real" information content in the data and for modelling the form of the depolarisation, as discussed in section 3.1 above, since now there are only two amplitudes to be determined, besides the parameters describing the model for $P_x(t)$. It is also possible to combine the two histograms with orthogonal phases to give a signal proportional to $P_x(t)$ directly, without any oscillating component.

3.3 Transforming data into the frequency domain

Case (b) If the muon sees a distribution of internal fields the counts in each detector will be:

$$C_j\,(t_n) = N_j^0\,\Delta t\,\exp(-\,t_n / \tau_\mu)\,[\,1 + A_j\,F_j(t_n)] \qquad (10)$$

where $$F_j(t_n) = \int_0^\infty F(\omega)\cos(\omega t_n - \phi_j)\,d\omega.$$

The function $F(\omega)$ gives the distribution of precession frequencies, which maps directly onto the distribution of internal fields through the gyromagnetic ratio: $\omega = \gamma B$. The problem is how to extract $F(\omega)$ from the data. This is clearly a Fourier transform problem, but standard Fourier methods have some drawbacks, in particular:

(i) The poor statistical accuracy of the µSR data at long times leads to very noisy transforms. Of course there are standard techniques for coping with this, for example use of window functions ("apodisation"). There is a large literature on how to choose the optimum window function, however whichever one is chosen, they all involve throwing away data, and some deterioration of the frequency resolution.

(ii) We have a problem of how to derive a unique frequency spectrum $F(\omega)$ from 32 separate histograms, with different phases ϕ_j.

Recently this problem has been tackled using the Maximum Entropy method [1]. This technique has been widely used to deal with a range of inverse problems in fields such as radio astronomy, geophysics and image reconstruction (including processing images from the Hubble telescope to correct for the aberrations in the mirror) etc. Applied to the present problem it has many advantages:

- It uses all the data available, but allows a unique frequency spectrum to be determined from multiple histograms.
- The frequency spectrum derived is necessarily positive.
- There are no prior assumptions about the form of $F(\omega)$.
- It gives the most uniform (N.B. *not* smoothest) distribution consistent with the data.
- It is possible to include convolution with the muon pulse shape.
- The ultimate frequency resolution can be achieved.

Details of the method will not be given here, since they are discussed in ref.[1]. Briefly the entropy S is defined in the information theoretic sense as

$$S = \sum_k (p_k / b_k)\log(p_k / b_k)$$

where the $\{p_k\}$ represent the heights of the bins in the histogram representing the frequency spectrum. The b_k represent the default values which give a scale to the magnitudes of the p_k. In practice the default is taken to be a flat level, independent of k. Starting from this featureless default level, the ME algorithm searches for a solution which maximises $S - \lambda\chi^2$, where λ is a Lagrange multiplier. Initially λ is chosen so that the algorithm puts most effort into minimising χ^2, with the aim of reducing its value to be equal to the number of data points. In the later stages the value of λ is altered to concentrate on maximising the entropy. The advantage of the technique is that it tackles the forward problem, i.e. generates a frequency spectrum which is then transformed, as in (10), to give the measured counts in each histogram, which are then compared to the measured values. This facilitates the incorporation of any transformation between the frequency domain and the time domain, e.g. convolution with the pulse shape of the muon bursts. The resulting frequency spectra produced by the method are impressive. For example Fig.1 shows the frequency spectrum for silver, which is very close to a delta function, as we might expect. This technique has been widely used by a number of groups [(2)] to study the mixed state of super- conductors. The distribution of internal fields in this case is highly asymmetric, with a high field tail extending up to B_{c2}, resulting from the flux lattice structure. The asymmetry of the μSR frequency distribution can therefore be used as an indicator of a well ordered flux lattice.

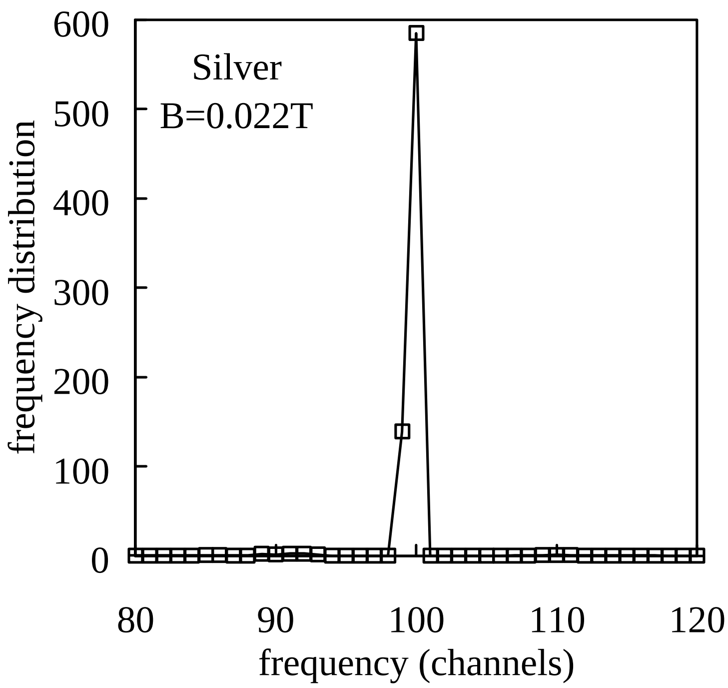

Figure 1. Frequency spectrum from a Maxent analysis of a silver data set with B_{ext}= 220 gauss. The width of one frequency channel corresponds to 30.5 kHz or 2.25 gauss. The small feature to the left of the main peak near channel 92 is due to muons stopping in the end of a collimator.

4. Limitations to the transverse μSR method.

4.1 Effect of finite muon lifetime.

Clearly μSR would not work if the muon did not decay! However the finite lifetime of the muon means that we can only collect useful data for, perhaps, five muon lifetimes (data at longer times are useful if we are prepared to count for a long time to improve the statistics). This immediately implies a limit to the frequency resolution of the transverse μSR technique. To see this consider the standard approach from the theory of discrete Fourier transforms (DFT): in a DFT of data comprising n channels with channel width Δt the frequency resolution Δf is simply $1/n\,\Delta t = 1/T$, i.e. the reciprocal of the total time for which data has been collected. If we take $T = 5\tau_\mu = 11$ μsec, then $\Delta f \approx 90$ kHz. Transforming this from frequency to magnetic field, using the gyromagnetic ratio of the muon, the resolution in magnetic field is about 0.7 mT or 7 gauss. Note that this is a fundamental restriction, no matter how good our instrumentation or data analysis.

4.2 Effect of finite pulse width.

The time structure of the muon beam at ISIS reflects that of the proton pulses at the production target, i.e. two pulses of width 70 ns (FWHM) with a separation of 340 ns. The repetition rate is 50 Hz. The pulses are further smeared by the 26 ns lifetime of the parent pions. The first of the double pulses is split between the two μSR instruments EMU and DEVA, while the second pulse passes on to the MUSR instrument. It follows that the time dependent count-rate (3) or (10) should be convoluted with the pulse shape of the incident muon beam. This will have a significant effect only when there are variations in the count-rate on a time scale comparable to the pulse width. There are two particular cases where this might be important: (i) when the Larmor precession frequency is high, and (ii) where there is rapid decay of the muon polarisation. Let us consider these in turn.

4.2 (i) Upper limit to the external field.

Suppose that we take the muon pulse shape to be a gaussian: $W(t) = \exp(-t^2/\tau_w^2)$. This is not strictly correct: a better approximation at ISIS is an inverted parabola, describing the proton pulse, convoluted with a decaying exponential, in order to account for the pion lifetime component. However a simple gaussian simplifies the result greatly. If we are studying a material like silver which is weakly damped, the resulting signal would be the convolution of the "ideal" signal $[1 + A\cos(\omega_L t)]$ with $W(t)$. Here A is the ideal asymmetry and ω_L is the Larmor precession frequency, as usual. It is straightforward to show that the effect of the convolution is to reduce the asymmetry to a value $A\exp[-(\omega_L\tau)^2/4]$. The reduction in the asymmetry with applied field is shown in Fig.2 for different values of the pulse width τ_w. It follows that the asymmetry falls to 1/e of its ideal value at an applied field B given by $(\omega_L/2) = (\gamma/2)\,B_{1/e} = 1/(\tau_w)$. For a pulse width of 70 ns (FWHM) the value of $\tau_w \approx 42$ ns, and the corresponding field $B_{1/e}$ is 560 gauss. The reduced asymmetry resulting from high Larmor precession rates degrades the signal to noise, as is clear from the discussion in section 1 above.

4.2 (ii) Upper limit to damping rate.

If the damping rate were very high then the muon polarisation $P_x(t)$ would fall to a small value within one Larmor precession period, or in the extreme case, within the muon pulse width. In the latter case, Lorentzian damping would only be visible for values of the damping rate smaller than λ_{max} given by, say, $\exp(-\lambda_{max}\tau_w) \approx e^{-1}$, that is we suppose that the muon polarisation decays by 1/e over the duration of the muon pulse. Using the above value $\tau_w = 42$ ns gives a value for $\lambda_{max} \approx 24$ μs^{-1}. However to extract such a value from real data would require a very accurate knowledge of the pulse shape and the start times for each histogram, and a careful deconvolution procedure. Notice that the situation at ISIS for measuring heavily damped signals is very much improved now that single pulses are used: with the same criterion as above for a double pulse of muons separated by 320 ns we find a value of λ_{max} of only 2.9μs^{-1}. In practice deconvolution of the pulse shape is necessary to extract reliable values of λ for damping rates greater than 3-5μs^{-1}, but the extra data analysis effort required is well justified.

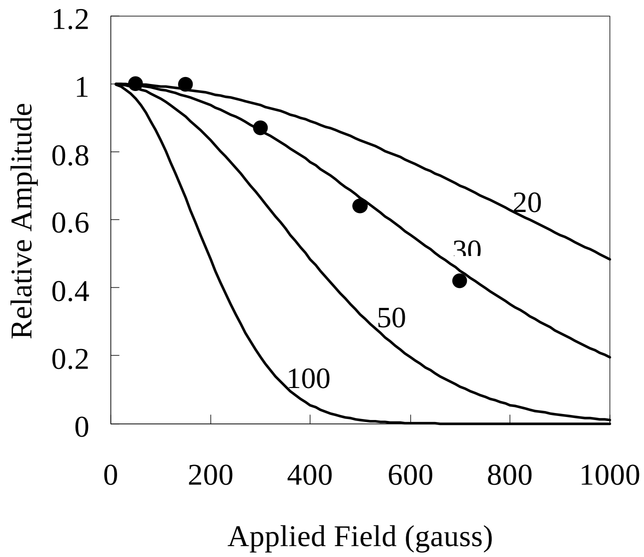

Figure 2 Variation of the asymmetry with external field due to the finite width of the muon pulses, after Nagamine and Yamakazi. The numbers on the curves give the value of τ_w in nanoseconds.

References

[1] B.D.Rainford and G.J.Daniell, Hyperfine Interactions **87** (1994) 1129.

[2] S.L.Lee et al., Phys. Rev. Lett. **71** (1993) 3862