Some Advance Methods of Statistical Analysis for the Muon g-2 Experiment at BNL

S. I. Redin*

Department of Physics, Yale University, New Haven, CT 06511
and Budker Institute of Nuclear Physics, Novosibirsk, Russia

Abstract. For the muon g-2 experiment at BNL a simple procedure was developed for analytical evaluation of statistical errors and correlations of parameters for the fit of time distribution of decay electrons with 5-parameter function $n(t) = N_\circ e^{-t/\tau}[1+A\cos(\omega t+\phi)]$. From 5 parameters $(N_\circ, \tau, A, \omega, \phi)$ the angular frequency of g-2 oscillations ω is the most important since it is one of two numbers (along with the magnetic field) to be measured precisely (at sub ppm level) in this experiment. It was shown that parameters ω and ϕ are correlated and knowledge of the phase of g-2 oscillations ϕ at the moment of injection of the muon beam into the storage ring can be used to improve the statistical accuracy of frequency ω. An appropriate formula for that case was derived. Methods of statistical analysis were also used to find systematic shifts of parameters of the 5-parameters function in case a small background of arbitrary time dependence is present.

The procedure, developed for statistical analysis for the muon g-2 experiment, was applied for statistical analysis of Breit-Wigner (nonrelativistic and relativistic) and Gaussian distributions.

A simple method to search for presence of small periodic background with known period is also presented. This method was developed for monitoring of the so called AGS flashlets background for the muon g-2 experiment at BNL and presently is routinely being used. It's based on "folding" of the time distribution of decay positrons into one AGS period of 2.694 μsec. For such applications this method supersedes the Fourier analysis method.

INTRODUCTION

One of the most common procedures of data analysis in high energy physics is fitting a histogram or distribution $n_i = n(t_i)$ with a function $f_{\vec{X}}(t)$ with one or several parameters $\vec{X}$, which are usually related to the values being studied in a given experiment. The main objective of the fitting is to find the best estimates for these parameters as well as their statistical errors and correlations. Although statistical errors and correlations come up routinely as an output of MINUIT or ROOT optimization, it is worthwhile to have the ability to estimate them independently of the fit itself. In this presentation we develop a simple procedure to derive a formula for statistical errors and correlations of parameters of a fit as a function of these parameters and statistics (usually $1/\sqrt{N}$ is involved). The procedure includes some integrals and for simple distributions analytical expressions can be obtained.

As a separate issue, in the last section we discuss a simple method, adopted in the muon g-2 experiment, to search for the presence of small periodic background with known period. This method has a clear advantage over Fourier analysis method for such an application.

STATISTICAL ERRORS FOR CORRELATED PARAMETERS

The general procedure for analytical evaluation of statistical errors and correlations can be better understood from the simplest case of 2 correlated parameters, say x and y. Let's assume a chi-square function $W(x,y)$ of a fit with a 2-parameter function has a minimum at $x_\circ$, $y_\circ$. Then the equation $\Delta W = W(x,y) - W(x_\circ, y_\circ) = A\,(\delta x)^2 + B\,\delta x\,\delta y + C\,(\delta y)^2 = 1$ defines the ellipse of errors, see Fig.1 from (1). Here $\delta x = x - x_\circ$, $\delta y = y - y_\circ$. From the definition of statistical errors σ_x, σ_y, also shown in Fig.1, one can derive $\sigma_x^{-2} = A - \frac{B^2}{4C}$, $\sigma_y^{-2} = C - \frac{B^2}{4A}$ (ABC formula).

Equation for multidimensional ellipsoid of errors for a fit with a multiparameter function $f_{\vec{X}}(t)$ would be

$$1 = \Delta W = \sum \frac{[f_{\vec{X}+\delta\vec{X}} - n_i]^2}{\sigma_i^2} = \sum \frac{[f_{\vec{X}+\delta\vec{X}} - f_{\vec{X}}]^2}{n_i} = \qquad (1)$$

* on behalf of g-2 collaboration

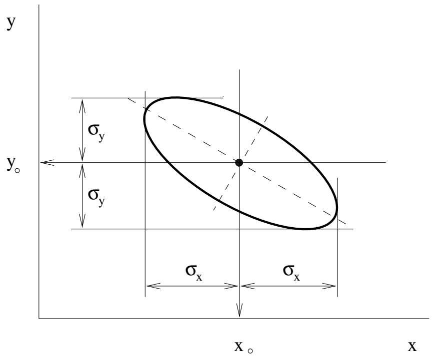

FIGURE 1. Ellipse of errors

$$= \sum \frac{\left(\sum_j \frac{\partial f}{\partial X_j}\delta X_j\right)^2}{f} \to \frac{N}{\int f dt}\int \frac{dt}{f}\left(\sum_j \frac{\partial f}{\partial X_j}\delta X_j\right)^2 = 1$$

In the last step we replace the sum by the integral: $\sum \to \int \frac{dt}{bin}$, where *bin* is the bin size of the histogram, and then get rid of that *bin* by using the number of events: $N = \sum n_i \to \frac{1}{bin}\int f_{\vec{X}}(t)\,dt$.

As an example and check of the procedure, let's find the errors of mean $t_\circ$, σ and amplitude A of a Gaussian distribution $f(t) = A\,e^{-\frac{(t-t_\circ)^2}{2\sigma^2}}$. From eq.1 find

$$1 = \Delta W = \frac{N}{\int f dt}\int \frac{dt}{f}\left(\frac{\partial f}{\partial t_\circ}\delta t_\circ + \frac{\partial f}{\partial \sigma}\delta\sigma + \frac{\partial f}{\partial A}\delta A\right)^2 =$$

$$= \frac{N}{\sigma^2}(\delta t_\circ)^2 + \frac{3N}{\sigma^2}(\delta\sigma)^2 + \frac{2N}{A\sigma}\delta\sigma\delta A + \frac{N}{A^2}(\delta A)^2 \qquad (2)$$

and then use the *ABC* formula to find $\sigma_{t_\circ} = \frac{\sigma}{\sqrt{N}}$, $\sigma_\sigma = \frac{\sigma}{\sqrt{2N}}$ and $\sigma_A = \sqrt{3}\frac{A}{\sqrt{2N}}$. The formers two equations are well known, of course (1). The procedure gives the correct result.

In this case only one cross term, namely $\delta A\delta\sigma$, does not vanish and hence A and σ are correlated and $t_\circ$ is statistically independent of A, σ. If, by any reason, parameter A is known and can be excluded from the set of fitted parameters, in eq.2 all terms with δA drop. As a result, $\sigma_{t_\circ}$ does not change but σ_σ does: $\sigma_\sigma = \frac{\sigma}{\sqrt{3N}} < \frac{\sigma}{\sqrt{2N}}$. In general, any additional information on one parameter of optimization improves the statistical precision of those parameters which correlate with this one and does not affect the others.

TIME DISTRIBUTION OF DECAY ELECTRONS IN MUON (G-2) EXPERIMENT

In the muon (g-2) experiment at BNL (2) the time distribution of decay electrons is fitted with a 5–parameter function $f(t) = N_\circ e^{-t/\tau}\left[1 + A\cos(\omega t + \phi)\right]$ in order to get precise value for the g-2 angular frequency ω. Other parameters are: amplitude A and phase ϕ of g-2 oscillations, muon lifetime in the storage ring, $\tau \approx 64.4\mu$sec, and normalization factor $N_\circ$. From the 5 parameters $N_\circ$ correlates with τ, and ω correlates with ϕ. For ω and ϕ exclusively, eq.1 gives

$$\frac{NA^2}{2}\left[(T^2 + 2T\tau + 2\tau^2)(\delta\omega)^2 + 2(T+\tau)\,\delta\omega\,\delta\phi + (\delta\phi)^2\right] = 1, \qquad (3)$$

where T is a histogram starting time. From eq.3 and the *ABC* formula find $\sigma_\phi = \frac{\sqrt{2}\sqrt{(T+\tau)^2+1}}{\tau A\sqrt{N}}$, $\sigma_\omega = \frac{\sqrt{2}}{\tau A\sqrt{N}}$. The latter can be used for various estimations and, in particular, for choosing an energy threshold E_{thr} of the decay electrons to maximize $A\sqrt{N}$ (both A and N are functions of E_{thr}). The errors of other parameters are: $\sigma_A = \frac{\sqrt{2}}{\sqrt{N}}$, $\sigma_\tau = \frac{\tau}{\sqrt{N}}$ and $\sigma_{N_\circ} = \frac{N_\circ}{\sqrt{N}}\sqrt{(T/\tau+1)^2+1}$.

It was known for years that the (g-2) frequency and phase correlate with each other, but only recently it was found "experimentally" that the magnitude of the correlation depends on the choice of time origin. Fig.2 shows result of fit for frequency and phase for the same histogram, first for regular case when $T \approx 70\mu$sec (ω and ϕ) and then for shifted time origin such that $T \approx -65\mu$sec (ω and ϕ'). The frequency does not change at all, but for the phase obvious correlations with frequency get greatly reduced (ϕ and ϕ' are in the same scale in Fig.2). That can be understood from eq.3: when $T\!=\!-\tau = -64.4\mu$sec, the $\delta\omega\,\delta\phi$ cross-term vanishes and so the $\omega\!-\!\phi$ (and $\tau\!-\!N_\circ$) correlations.

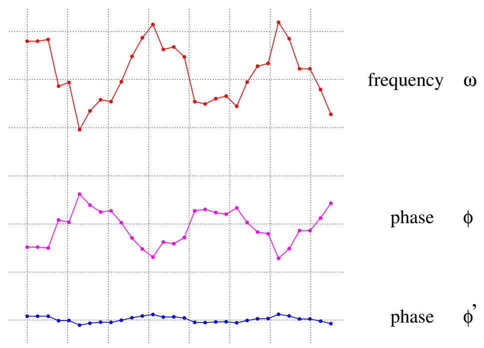

FIGURE 2.

In the muon (g−2) experiment, a beam of polarized muons is injected into the storage ring and at that point the muon polarization vector is directed at a certain small anlge to the direction of motion. Thus, in general, the phase of the (g−2) oscillations is known at the injection time and one can use this knowledge to improve the precision of (g-2) frequency ω. In a special case when the phase is known at some arbitrary time T' with precision σ_F we should add a term $\frac{[\delta(\omega T'+\phi)]^2}{\sigma_F^2} = \sigma_F^{-2}\,[T'^2\,(\delta\omega)^2 + 2T'\,(\delta\omega)(\delta\phi) + (\delta\phi)^2]$ in eq.3. Then after little algebra we find

$$\sigma_\omega^{-2} = \sigma_{\omega\circ}^{-2}\left[1 + \frac{\sigma_F^{-2}}{(\tau\sigma_{\omega\circ})^{-2} + \sigma_F^{-2}}\,\frac{(T+\tau-T')^2}{\tau^2}\right], \qquad (4)$$

where $\sigma_{\omega\circ} = \frac{\sqrt{2}}{\tau A\sqrt{N}}$.

SYSTEMATIC SHIFT DUE TO SMALL BACKGROUND

A procedure, similar to that developed above for the statistical errors and correlations, can be developed for estimation of the systematic shift of fit parameters in case a small background is present in the data. Suppose we have some low level background $h(t)$ admixed to the data, otherwise perfectly described by some multi-parameter function $f_{\vec{X}}(t)$. The background is small enough to observe "by eyes" or even to spoil the chi-squared considerably. Nevertheless fitting the histogram with a function $f_{\vec{X}}(t)$ alone will give us parameters $\vec{X}$, shifted with respect to the "true" parameters $\vec{X}_\circ$: $\vec{X} = \vec{X}_\circ + \delta\vec{X}$.

We can find the systematic shift $\delta\vec{X}$ from the chi-squared minimization requirement $\partial W/\partial X_j = 0$:

$$\frac{\partial W}{\partial X_j} = \frac{\partial}{\partial X_j}\sum_i \frac{\left(f_{\vec{X}_\circ+\delta\vec{X}} - n_i\right)^2}{\sigma_i^2} = 2\sum_i \frac{\left(f_{\vec{X}_\circ+\delta\vec{X}} - n_i\right)}{n_i}\,\frac{\partial f}{\partial X_j} =$$

$$= 2\sum_i \frac{\left(f_{\vec{X}_\circ+\delta\vec{X}} - f_{\vec{X}} - h\right)}{f_{\vec{X}}}\,\frac{\partial f}{\partial X_j} = 2\sum_i \frac{\left(\sum_k \frac{\partial f}{\partial X_k}\delta X_k - h\right)}{f}\,\frac{\partial f}{\partial X_j} =$$

$$= 2\sum_k \delta X_k \sum_i \frac{1}{f}\left(\frac{\partial f}{\partial X_k}\,\frac{\partial f}{\partial X_j}\right) - 2\sum_i \frac{h}{f}\frac{\partial f}{\partial X_j} = 0 \qquad (5)$$

Replacing the sum by an integral we finally get a system of linear equations:

$$\sum_k \delta X_k \int \frac{dt}{f}\,\frac{\partial f}{\partial X_k}\,\frac{\partial f}{\partial X_j} = \int dt\,\frac{h}{f}\frac{\partial f}{\partial X_j} \qquad (6)$$

For g-2 distribution $N_\circ e^{-t/\tau}[1 + A\cos(\omega t+\phi)]$ eq.6 gives systematic shifts

$$\begin{aligned}
\delta N_\circ &= \frac{1}{e\tau}\int_{-\tau}^{\infty} h(t)\,dt \\
\delta\tau &= \frac{1}{eN_\circ\tau}\int_{-\tau}^{\infty} t\,h(t)\,dt \\
\delta A &= \frac{2}{eN_\circ\tau}\int_{-\tau}^{\infty}\frac{h(t)\cos(\omega t+\phi)}{1+A\cos(\omega t+\phi)}\,dt \qquad (7)\\
\delta\phi &= -\frac{2}{eN_\circ A\tau}\int_{-\tau}^{\infty}\frac{h(t)\sin(\omega t+\phi)}{1+A\cos(\omega t+\phi)}\,dt
\end{aligned}$$

$$\begin{aligned}
\delta\omega = &-\frac{2}{eN_\circ A\tau^3}\int_{-\tau}^{\infty}\frac{t\,h(t)\sin(\omega t+\phi)}{1+A\cos(\omega t+\phi)}\,dt + \\
&+\frac{2}{e^2N_\circ A\tau^4}\int_{-\tau}^{\infty}h(t)\,dt\int_{-\tau}^{\infty}t\,e^{-t/\tau}\sin(\omega t+\phi)\,dt + \\
&+\frac{2}{e^2N_\circ A\tau^6}\int_{-\tau}^{\infty}t\,h(t)\,dt\int_{-\tau}^{\infty}t^2\,e^{-t/\tau}\sin(\omega t+\phi)\,dt
\end{aligned}$$

For $\delta\omega$ we have also included two next-to-leading terms, which are important if $h(t)$ is a smooth function. Note that we chose the time origin in such a way that the histogram starting time is $T = -\tau$. That doesn't change result, only makes the calculations easier.

BREIT−WIGNER DISTRIBUTION

In approximation of narrow width the nonrelativistic Breit−Wigner distribution can be written as $f(E) = A\,\frac{\Gamma^2/4}{(E-M)^2+\Gamma^2/4}$. For statistical errors of parameters M, Γ and A eq.1 gives

$$\frac{2N}{\Gamma^2}(\delta M)^2 + \frac{3N}{2\Gamma^2}(\delta\Gamma)^2 + \frac{2N}{A\Gamma}\delta\Gamma\,\delta A + \frac{N}{A^2}(\delta A)^2 = 1 \qquad (8)$$

and from that

$$\sigma_M = \frac{\Gamma}{\sqrt{2N}},\quad \sigma_\Gamma = \sqrt{2}\,\frac{\Gamma}{\sqrt{N}},\quad \sigma_A = \sqrt{3}\,\frac{A}{\sqrt{N}} \qquad (9)$$

As follows from eq.8, parameters Γ and A are correlated and M is independent.

For the Breit−Wigner plus constant background distribution $f(E) = A\,\frac{\Gamma^2/4}{(E-M)^2+\Gamma^2/4} + C$ parameter M is still independent but the rest three parameters Γ, A and C correlate with each other. That makes the algebra calculation much more tedious. The result is

$$\begin{aligned}
\sigma_M &= \frac{\Gamma}{\sqrt{2N}}\sqrt{\frac{C}{A}}\sqrt{\frac{(1+Z)^3}{3+Z}} \\
\sigma_\Gamma &= \sqrt{2}\frac{\Gamma}{\sqrt{N}}\sqrt{\frac{C}{A}}\,(1+Z) \qquad (10)\\
\sigma_A &= \frac{A}{\sqrt{N}}\sqrt{\frac{C}{A}}\,\sqrt{Z(1+3Z)} \\
\sigma_C &= \frac{C}{\sqrt{N_c}}
\end{aligned}$$

where $N \propto A\Gamma\pi/2$ is a number of Breit−Wigner events, $N_c \propto (E_{max} - E_{min})C$ is a number of background events and $Z = \sqrt{1 + A/C}$.

The formula for the relativistic Breit−Wigner distribution $f(s) = A\,\frac{m^2\Gamma^2}{(s-m^2)^2+m^2\Gamma^2}$ can be written as $f(s) = A\,\frac{Y^2/4}{(s-X)^2+Y^2/4}$ with $X = m^2$ and $Y = 2m\Gamma$ and we can use eq.9: $\sigma_X = \frac{\Gamma}{\sqrt{2N}}$, $\sigma_Y = \sqrt{2}\,\frac{Y}{\sqrt{N}}$, $\sigma_A = \sqrt{3}\,\frac{A}{\sqrt{N}}$. Making inverse transformation, find statistical errors for the relativistic Breit−Wigner distribution: $\sigma_m = \frac{\Gamma}{\sqrt{2N}}$, $\sigma_\Gamma =$

$\sqrt{2}\,\frac{\Gamma}{\sqrt{N}}\sqrt{1+\frac{\Gamma^2}{4m^2}} \approx \sqrt{2}\,\frac{\Gamma}{\sqrt{N}}$, $\sigma_A = \sqrt{3}\,\frac{A}{\sqrt{N}}$, which are exactly same as for the nonrelativistic case, see eq.9.

FOLDING METHOD TO MONITOR THE AGS BACKGROUND

In the first run of the muon (g−2) experiment at BNL in 1997 an accelerator background admixture was found in (g−2) data. It was caused by a proton leakage after extraction of the main pulse out of the AGS into our beam line. It revealed itself as a fence-like structure over the regular (g−2) wiggles, see Fig.3. The time intervals between these “flaslets” are 2.7 μsec, i.e. equal to the revolution time for a proton bunch in the AGS, T_{AGS}. That was the main reason to associate these to an AGS proton leakage. The problem was discussed with the AGS experts and various schemes of solution, including fast shutter magnet in the (g−2) beam line, were proposed and implemented. Still, the monitoring for the presence of the AGS flaslet background was and is very important.

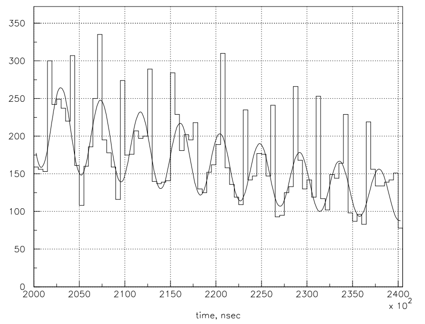

FIGURE 3. Fragment of decay positrons time distribution for the run #883

A very sensitive method to monitior the AGS background was developed. The idea is to put all the data into one AGS period of 2.7 μsec. To do that, the transformation

$$t' = \frac{t}{T_{AGS}} - \text{integer}\left(\frac{t}{T_{AGS}}\right) \qquad (11)$$

was applied to the data. Fig. 4 shows the result from one of the longest 1997 runs (#1041) without AGS background. The (g–2) oscillations average out completely. The exponential decay with time constant $\tau = \gamma \cdot \tau_\mu = 64.4\mu$sec appears as a 4% slope, as shown by the solid line. The distribution has a maximum at t'=0.15 which corresponds to the start of data taking.

The distribution of t' for the run #883 is shown in Fig. 5. The highest peak at t' = 0.90 corresponds to the 2.7 μsec structure seen in Fig. 3. Also 7 smaller peaks are clearly seen. All together the 8 peaks correspond to the 8 bunches in the AGS. The distribution in t' is very sensitive

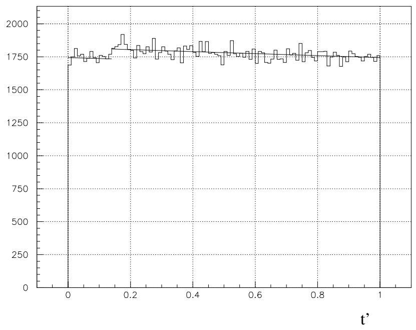

FIGURE 4. Distribution of t' for the run #1041

to the exact value of the AGS period T_{AGS}. Comparing plots for several T_{AGS} with 1 nsec interval, we clearly saw the sharpest structure for $T_{AGS} = 2.694\,\mu$sec. AGS experts confirmed that 2.694 μsec is indeed the best estimate for the AGS period.

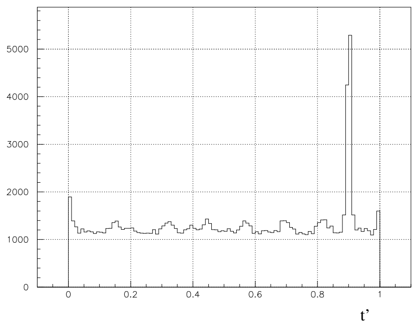

FIGURE 5. Distribution of t' for the run #883

For monitoring the AGS background the folding method (also known in the (g−2) collaboration as a “Redin plot”) has some advantages compared to the Fourier analysis method. Fourier analysis works well for sinewave structures. But for the AGS background shown in Fig. 3, which looks like a set of equally spaced delta–functions in time, the Fourier analysis would give a similar set of equally spaced delta–functions in a frequency domain, where the statistical analysis is complicated. Furthermore the 8 beam structure can be better understood from analysis of the t' distribution than from the Fourier analysis.

REFERENCES

1. Review of Particle Physics, Eur. Phys. J. **C15**, 1(1998)
2. R. M. Carey et al., Phys. Rev. Lett. **82**, 1632 (1999); H. N. Brown et al., Phys. Rev. **D62**, 091101 (2000)