Statistical errors and correlations of parameters in data analysis for the muon (g–2) experiment at BNL

S. I. Redin [a,b,1] et al.

[a] *Budker Institute of Nuclear Physics, Novosibirsk, Russia*

[b] *Physics Department, Yale University, New Haven, CT 06511, USA*

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 $G(t) = N_\circ \, e^{-t/\tau} \, [1 + A \, \cos(\omega \, t + \phi)]$. 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 might be used to improve the statistical accuracy of ω. An appropriate equation for that case is derived.

Some generic equations given in this paper can also be applied to other distributions in physics and elsewhere.

Key words: Statistical errors, Correlations, Fitting

1 Introduction

In an ongoing muon (g–2) experiment E821 at Brookhaven National Laboratory [1] two quantities must be measured precisely in order to obtain the muon (g–2) value to required accuracy. One of those is the muon spin precession frequency ω_a in uniform magnetic field and another is the average magnetic field in the muon storage volume, measured by advanced NMR probe system in terms of proton spin precession frequency ω_p.

[1] Corresponding author. E-mail sredin@sun2.bnl.gov

The muon spin precession frequency ω_a appears as a modulation of the time spectrum of muon decay electrons. Electrons from muon decay are detected with calorimeters and their times of arrival are accurately measured. The time distribution of electrons is basically an exponentially decaying sine wave with period being $T_a = 2\pi/\omega_a$. The data are fitted with an appropriate fit function and thus ω_a and its statistical error are obtained.

Fitting a histogram is a typical method of data analysis in high energy physics, however accuracy requierment for muon (g–2) experiment ($\sim 3\cdot 10^{-7}$) is exceptionally high. Therefore all statistical properties and correlations of parameters must be well understood. In this paper we analyse statistical properties of the five parameter distribution :

$$G(t) = N_\circ\, e^{-t/\tau} \left[\, 1 + A\, \cos(\omega_a t + \phi) \,\right] \tag{1}$$

which, with or without modifications, is of central importance for the muon (g–2) data analysis. The five parameters in eq.1 are: normalization factor $N_\circ$, lifetime τ and amplitude A, frequency ω_a and phase ϕ of g–2 oscillations. In the rest of the paper we omit subscript a at the muon g–2 frequency ω as it is the only frequency to be considered.

2 Statistical errors of fit parameters

In a typical data analysis procedure, parameters $x_1, x_2, ..., x_L$ of fit function $f(x_1, x_2, ..., x_L; t)$ are obtained by the least squares method, which requires minimization of χ^2 as a function of parameters x_i. For simplicity we denote $W \equiv \chi^2$, and vector of parameters $(x_1, x_2, ..., x_L)$ as $\boldsymbol{x}$. Then, by definition

$$W(\boldsymbol{x}) = \sum_n \frac{\left(f(\boldsymbol{x}; t_n) - \mathcal{N}_n\right)^2}{\sigma_n^2} \tag{2}$$

where t_n and $\mathcal{N}_n$ are center of n^{th} channel of the histogtam and its content, respectively, and σ_n is statistical error associated with n^{th} channel. The latter is often taken as $\sigma_n^2 = \mathcal{N}_n$ as been good approximation for large $\mathcal{N}_n$ and simple technically. However, for muon (g–2) data analysis it was proved to be not accurate enough and more rigorous expression $\sigma_n^2 = f(\boldsymbol{x}; t_n)$ is always used. Thus

$$W(\boldsymbol{x}) = \sum_n \frac{\left(f(\boldsymbol{x}; t_n) - \mathcal{N}_n\right)^2}{f(\boldsymbol{x}; t_n)} \tag{3}$$

One might find parameters of optimization by requiring derivatives of $W(\boldsymbol{x})$ with respect to each parameter to vanish: $\partial W/\partial x_i = 0,\ i = 1, 2, ..., L$. That would lead to a system of L equations (in general nonlinear) for L variables x_i. Such a system is usually been solved numerically by MINUIT [2] or other standard minimization routines with great accuracy.

Statistical errors of parameters are obtained from analysis of $W(\boldsymbol{x})$ in vicinity of point of optimum $\boldsymbol{x}_\circ \equiv (x_{1\circ}, x_{2\circ}, ..., x_{L\circ})$. $W(\boldsymbol{x})$ can be written as

$$W(\boldsymbol{x}) = W(\boldsymbol{x}_\circ) + \Delta W(\delta\boldsymbol{x}) \tag{4}$$

where $\delta\boldsymbol{x} = \boldsymbol{x} - \boldsymbol{x}_\circ$ and $\Delta W(0) = 0$. Equation

$$\Delta W(\delta\boldsymbol{x}) = 1\,, \tag{5}$$

which defines a shell in the $L-$dimensional space of parameters x_i or δx_i, is used to define statistical errors σ_i as an extreme extends of the shell (5) in a given dimension (or direction):

$$\sigma_i^+ = \max\,\delta x_i\,, \qquad \sigma_i^- = |\,\min\,\delta x_i\,| \tag{6}$$

MINUIT, etc. find these statistical errors numerically fairly precisely in general case. In a simple case or in a reasonable approximation, eq.(5) becomes particular simple:

$$\Delta W = \sum_{i=1}^{L}\sum_{j=1}^{L} a_{ij}\ \delta x_i\ \delta x_j = 1 \tag{7}$$

with $a_{ij} = a_{ji}$ for simplicity. Terms linear in δx_i are absent in eq.7 by definition of minimum of function $W(\boldsymbol{x})$. Eq.7 describes a perfect ellipsoid in a $L-$dimensional space (given W have minimum at $\delta\boldsymbol{x} = 0$). In such a case, as proved in Appendix I, the errors can be found analytically as

$$\sigma_i = \sqrt{\frac{M_{ii}}{det(\mathcal{A})}} \tag{8}$$

where $det(\mathcal{A})$ is determinant of matrix of coefficients a_{ij} (error matrix) in eq.(7) and M_{ii} is its $ii^{\,th}$ (i.e. diagonal) minor. For smooth function $f(\boldsymbol{x};t)$ and large $\mathcal{N}_n$, a_{ij} can be found analytically in a reasonably good approximation:

$$1 = \Delta W = \sum_n \frac{(f - \mathcal{N}_n)^2}{f} - \sum_n \frac{(f_\circ - \mathcal{N}_n)^2}{f_\circ} \approx \sum_n \frac{(f - \mathcal{N}_n)^2 - (f_\circ - \mathcal{N}_n)^2}{f_\circ} =$$

$$= \sum_n \frac{f + f_\circ - 2\mathcal{N}_n}{f_\circ}(f - f_\circ) \approx \sum_n \frac{(f - f_\circ)^2}{f_\circ} \approx \sum_n \left(\sum_{i=1}^{L} \sum_{j=1}^{L} \frac{f'_i \, f'_j}{f_\circ} \delta x_i \, \delta x_j \right) =$$

$$= \sum_{i=1}^{L} \sum_{j=1}^{L} \left(\sum_n \frac{f'_i \, f'_j}{f_\circ} \right) \delta x_i \, \delta x_j \ , \qquad \text{and hence} \quad a_{ij} = \sum_n \frac{f'_i \, f'_j}{f_\circ} \tag{9}$$

Here $f'_i = (\partial f / \partial x_i)_{\boldsymbol{x}=\boldsymbol{x}_\circ}$ and $f_\circ = f(\boldsymbol{x}_\circ; t)$. For histogram bin width b significantly shorter than typical structure of a distribution, e.g. period of muon (g–2) oscillations in our case, summation in eq.9 can be replaced by integration: $\sum_n \to \int dt/b$. In turn, parameter b can be replaced by total number of events N through the normalization relation:

$$N = \sum_n \mathcal{N}_n \approx \sum_n f_\circ(t_n) \approx \int f_\circ(t) \frac{dt}{b} \qquad \text{hence} \quad \frac{1}{b} = \frac{N}{\int f_\circ(t) \ dt} \tag{10}$$

Finally:

$$a_{ij} = \frac{N}{\int f_\circ(t) \ dt} \quad \int \frac{f'_i \, f'_j}{f_\circ} dt \tag{11}$$

with explicitly symmetrical coefficients $a_{ij} = a_{ji}$. Since certain approximations applied, analytical expressions, based on eqs.8 and 11, in general are not that precise as numerical calculations are. Nonetheless they are extremely useful as revealing, in explicitly functional form, statistical relations among parameters of the problem. That was proved to help deeper understanding of nature of the problem and stimulate further improvements of the data analysis and experiment as a whole. This is certainly true for our g–2 data analysis.

3 Statistical errors of parameters of the 5-parameter fit

Let us calculate statistical errors of the 5-parameter fit (eq.1) used for muon (g–2) data analysis. Find derivatives of $G(t) = N_\circ \, e^{-t/\tau} [\, 1 + A \, \cos(\omega t + \phi) \,]$ with respect to its parameters:

$$f'_1 = \frac{\partial G}{\partial N_\circ} = e^{-t/\tau} \, [\, 1 + A \, \cos(\omega t + \phi) \,] \tag{12}$$

$$f'_2 = \frac{\partial G}{\partial \tau} = N_\circ \, \frac{t}{\tau^2} \, e^{-t/\tau} \, [\, 1 + A \, \cos(\omega t + \phi) \,] \tag{13}$$

$$f'_3 = \frac{\partial G}{\partial A} = N_\circ \, e^{-t/\tau} \, \cos(\omega t + \phi) \tag{14}$$

$$f'_4 = \frac{\partial G}{\partial \omega} = -N_\circ \, e^{-t/\tau} \, A \, t \, \sin(\omega t + \phi) \tag{15}$$

$$f_5' = \frac{\partial G}{\partial \omega} = -N_\circ \, e^{-t/\tau} \, A \, \sin(\omega t + \phi) \tag{16}$$

Calculation of relevant integrals in eq.10,11 is given in Appendix II. In our evaluations we neglect integrals of fast oscillating functions like $t\, e^{-t/\tau} \sin(\omega t + \phi)$, those relative contribution is of order $\epsilon \equiv (\omega\tau)^{-1} = 0.01$. To make calculations of statistical errors in eq.8 easy, we can simplify matrix $\mathcal{A}(a_{ij})$ by replacing our 5-component vector of parameters $(\delta N_\circ \; \delta\tau \; \delta A \; \delta\omega \; \delta\phi) \equiv (x_1 \; x_2 \; x_3 \; x_4 \; x_5)$ by vector $(\frac{\sqrt{N}}{N_\circ} x_1 \;\; \frac{\sqrt{N}}{\tau} x_2 \;\; \sqrt{\frac{N}{2}}\, x_3 \;\; \sqrt{\frac{N}{2}} A\tau \, x_4 \;\; \sqrt{\frac{N}{2}} A \, x_5) \equiv (y_1 \; y_2 \; y_3 \; y_4 \; y_5)$. Then eq.7 in matrix form will read:

$$(y_1 \; y_2 \; y_3 \; y_4 \; y_5) \times \begin{pmatrix} 1 & \frac{T}{\tau}+1 & 0 & 0 & 0 \\ \frac{T}{\tau}+1 & (\frac{T}{\tau}+1)^2+1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & (\frac{T}{\tau}+1)^2+1 & \frac{T}{\tau}+1 \\ 0 & 0 & 0 & \frac{T}{\tau}+1 & 1 \end{pmatrix} \times \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{pmatrix} = 1 \tag{17}$$

Here T is a histogram starting time. Calculating determinant of 5×5 matrix in eq.20 and its diagonal minors, one finds:

$$\sigma_{y1} = \sigma_{y5} = \sqrt{(T/\tau + 1)^2 + 1}\,, \quad \sigma_{y2} = \sigma_{y3} = \sigma_{y4} = 1 \tag{18}$$

and errors of the g–2 5-parameter fit:

$$\begin{aligned} \sigma_{N_\circ} &= \frac{N_\circ}{\sqrt{N}}\, \sigma_{y1} = \frac{N_\circ}{\sqrt{N}} \sqrt{(T/\tau + 1)^2 + 1} \\ \sigma_\tau &= \frac{\tau}{\sqrt{N}}\, \sigma_{y2} = \frac{\tau}{\sqrt{N}} \\ \sigma_A &= \frac{\sqrt{2}}{\sqrt{N}}\, \sigma_{y3} = \frac{\sqrt{2}}{\sqrt{N}} \\ \sigma_\omega &= \frac{\sqrt{2}}{\tau A \sqrt{N}}\, \sigma_{y4} = \frac{\sqrt{2}}{\tau A \sqrt{N}} \\ \sigma_\phi &= \frac{\sqrt{2}}{A\sqrt{N}}\, \sigma_{y5} = \frac{\sqrt{2}}{A\sqrt{N}} \sqrt{(T/\tau + 1)^2 + 1} \end{aligned} \tag{19}$$

The error matrix $\mathcal{A}(a_{ij})$ in eq.20 is a diagonal-block matrix and hence eq.20 can be split into three independent matrix equations with lower rank error matrixes. That might simplify calculations substantially.

4 Correlations of parameters of the 5-parameter fit

The nonvanishing off-diagonal elements $a_{12} = a_{21} = T/\tau + 1$ and $a_{45} = a_{54} = T/\tau + 1$ of error matrix in eq.20 indicate the only correlations among parameters of the 5-parameter fit: $N_\circ$ and τ and ω and ϕ, respectively. These elements, and hence correlations, vanish when $T = -\tau$, i.e. when histogram starting time is happen to be at one lifetime τ prior to the time origin point $t = 0$. This is rather unexpected feature that such a fundamental value as correlation depends on the choice of time origin. Furthermore, as follows from eq.19, σ_τ and $\sigma_{N_\circ}$ also depend on T and hence on the time origin and both have minimum at $T = -\tau$. The reason for that is, in a sense, a hidden definition of ϕ and $N_\circ$: these are phase and magnitude of $G(t)$ at $t = 0$, respectively, and as such depend indeed on the choice of time origin. Note that in spite of correlation with ϕ, σ_ω does not depend on T as expected. This is also true for σ_τ. In this case error matrix $\mathcal{A}$ becomes diagonal and eq.7 in a matrix form can be written as

$$(\delta N_\circ\ \delta\tau\ \delta A\ \delta\omega\ \delta\phi) \times \begin{pmatrix} \frac{N}{N_\circ^2} & 0 & 0 & 0 & 0 \\ 0 & \frac{N}{\tau^2} & 0 & 0 & 0 \\ 0 & 0 & \frac{N}{2} & 0 & 0 \\ 0 & 0 & 0 & \frac{N\,A^2\tau^2}{2} & 0 \\ 0 & 0 & 0 & 0 & \frac{N\,A^2}{2} \end{pmatrix} \times \begin{pmatrix} \delta N_\circ \\ \delta\tau \\ \delta A \\ \delta\omega \\ \delta\phi \end{pmatrix} = 1 \qquad (20)$$

Since the time frame where $T = -\tau$ (we shall call such a frame as a center of gravity frame) is a special one, let us define phase ϕ in an arbitrary time frame through the phase in the center of gravity frame ϕ'. We denote t and t' to be time in an arbitrary and the center of gravity frames, respectively. Then $t' = t - (T + \tau)$ and $\omega t + \phi = \omega t' + \omega(T + \tau) + \phi = \omega t' + \phi'$, thus

$$\phi' = \omega(T + \tau) + \phi \quad \text{and} \quad \phi = \phi' - \omega(T + \tau) \qquad (21)$$

Since frequency and phase are statistically independent in the center of gravity frame (correlations vanish for $T = -\tau$), one can use eq.21 to evaluate statistical error of ϕ in an arbitrary time frame:

$$\sigma_\phi^2 = \sigma_{\phi'}^2 + \sigma_\omega^2 (T + \tau)^2 = \frac{2}{A^2 N} + \frac{2}{\tau^2 A^2 N}(T + \tau)^2 = \frac{2}{A^2 N}[(T/\tau + 1)^2 + 1]$$
$$\text{and} \qquad \sigma_\phi = \frac{\sqrt{2}}{A\sqrt{N}}\sqrt{(T/\tau + 1)^2 + 1} \qquad (22)$$

which is identical to eq.19. Same can be done for $\sigma_{N_\circ}$.

5 Possible improvement of σ_ω by knowledge of phase

Frequency and phase correlation is a base of possible improvement of statistical error σ_ω by knowledge of phase in the muon g–2 experiment. In the experiment beam of polarized muons is injected into the muon storage ring from the beam line. At the injection point spin of muons is parallel to their momentum in the case of positive muons and antiparallel in the case of positive muons. For simplicity we set time origin at the moment of injection and assume that muon polarization, and hence phase in eq.1, is known precisely. In this case ϕ can be excluded from the set of fit parameters. To calculate σ_ω for such a case, one can use eq.20 with erased y_5 in both vectors and 5^{th} row and column in matrix $\mathcal{A}$. The result is

$$\sigma_{y4} = \frac{1}{\sqrt{(T/\tau+1)^2+1}}\,, \quad \text{and hence} \quad \sigma_\omega = \frac{\sqrt{2}}{\tau\, A\,\sqrt{N}}\frac{1}{\sqrt{(T/\tau+1)^2+1}} \tag{23}$$

which is less than σ_ω in eq.19 by factor of $\sqrt{(T/\tau+1)^2+1}$. Since data taking always starts later than injection, $T \geq 0$ and $\sqrt{(T/\tau+1)^2+1} \geq \sqrt{2}$ and hence accuracy of ω can be improved by more than a factor $\sqrt{2}$ if phase (i.e. average polarization of muons) is known precisely at the moment of injection. In current muon (g–2) experiment data taking can be started not early than some 32 μs after injection due to scraping procedure mentioned in [1]. For $T/\tau = 32\,\mu\text{s}/64.4\,\mu\text{s} \approx 0.5$, precise knowledge of phase may give improvement as much as by factor of $\sqrt{(0.5+1)^2+1} = \sqrt{3.25} = 1.8$, which is equivalent to 3.25–fold increase of statistics or 3.25–fold decrease of required running time.

We now consider more realistic and general case when phase ϕ' is known at some time T' with finite precision σ_F independently of fit, and histogram starts at time T in an arbitrary time frame. We restrict our study by two variables ($\delta\omega\ \delta\phi$) since other parameters are statistically independent of these two. Additional constrain on phase would give an extra term $\Delta W'$:

$$\Delta W' = \frac{(\delta\phi')^2}{\sigma_F^2} = \frac{(\delta\omega\, T' + \delta\phi)^2}{\sigma_F^2} = \frac{T'^2}{\sigma_F^2}(\delta\omega)^2 + 2\frac{T'}{\sigma_F^2}\,\delta\omega\,\delta\phi + \frac{1}{\sigma_F^2}(\delta\phi)^2 \tag{24}$$

to ΔW (eqns.5 and 7) which, in turn, we rewrite as

$$\Delta W(\delta\omega,\delta\phi) = \frac{NA^2\tau^2}{2}\left[(T/\tau{+}1)^2{+}1\right](\delta\omega)^2 + NA^2\tau(T/\tau{+}1)\,\delta\omega\,\delta\phi + \frac{NA^2}{2}(\delta\phi)^2 =$$
$$= \sigma_{\omega\circ}^{-2}\left[(T/\tau{+}1)^2{+}1\right](\delta\omega)^2 + 2\frac{\sigma_{\omega\circ}^{-2}}{\tau}(T/\tau{+}1)\,\delta\omega\,\delta\phi + \frac{\sigma_{\omega\circ}^{-2}}{\tau^2}(\delta\phi)^2 \tag{25}$$

We use relations

$$\omega t + \phi = \omega(t - T') + \omega T' + \phi \qquad \text{and subsequently} \qquad \phi' = \omega T' + \phi \quad (26)$$

to express phase at time T' in terms of parameters of optimization ϕ and ω in eq.24. We also introduce $\sigma_{\omega\circ} = \frac{\sqrt{2}}{\tau A \sqrt{N}}$ which is statistical error of ω from the 5-parameter fit alone. Then total error matrix will read:

$$\mathcal{A} = \begin{pmatrix} \sigma_{\omega\circ}^{-2}\,[(\frac{T}{\tau}+1)^2+1] + \sigma_F^{-2}\,T'^2 & \sigma_{\omega\circ}^{-2}\frac{1}{\tau}(\frac{T}{\tau}+1) + \sigma_F^{-2}\,T' \\ \sigma_{\omega\circ}^{-2}\frac{1}{\tau}(\frac{T}{\tau}+1) + \sigma_F^{-2}\,T' & \sigma_{\omega\circ}^{-2}\frac{1}{\tau^2} + \sigma_F^{-2} \end{pmatrix} \quad (27)$$

and we want to use equation $\sigma_\omega^{-2} = det(\mathcal{A})/M_{11}$ to find statistical error of ω. Since determinant of error matrix $\mathcal{A}$ in eq.27 is equal to product of diagonal elements minus product of non-diagonal elements, and minor of the upper left element is simply equial to the lower right element, σ_ω^{-2} can be calculated as

$$\begin{aligned}
\sigma_\omega^{-2} &= \sigma_{\omega\circ}^{-2}\,\left[(T/\tau+1)^2+1\right] + \sigma_F^{-2}\,T'^2 - \frac{\left(\sigma_{\omega\circ}^{-2}\frac{1}{\tau}(\frac{T}{\tau}+1)+\sigma_F^{-2}\,T'\right)^2}{\sigma_{\omega\circ}^{-2}\frac{1}{\tau^2}+\sigma_F^{-2}} = \\
&= \sigma_{\omega\circ}^{-2} + \frac{\left(\sigma_{\omega\circ}^{-2}\,(\frac{T}{\tau}+1)^2+\sigma_F^{-2}\,T'^2\right)\left(\sigma_{\omega\circ}^{-2}\frac{1}{\tau^2}+\sigma_F^{-2}\right) - \left(\sigma_{\omega\circ}^{-2}\frac{1}{\tau}(\frac{T}{\tau}+1)+\sigma_F^{-2}\,T'\right)^2}{\sigma_{\omega\circ}^{-2}\frac{1}{\tau^2}+\sigma_F^{-2}} = \\
&= \sigma_{\omega\circ}^{-2} + \frac{\sigma_{\omega\circ}^{-2}\sigma_F^{-2}\left(\frac{T'^2}{\tau^2}+(\frac{T}{\tau}+1)^2-2\frac{T'}{\tau}(\frac{T}{\tau}+1)\right)}{\sigma_{\omega\circ}^{-2}\frac{1}{\tau^2}+\sigma_F^{-2}} = \qquad (28)\\
&= \sigma_{\omega\circ}^{-2}\left(1+\frac{\sigma_F^{-2}}{\sigma_{\omega\circ}^{-2}\frac{1}{\tau^2}+\sigma_F^{-2}}\frac{(T+\tau-T')^2}{\tau^2}\right) = \sigma_{\omega\circ}^{-2}\left(1+\frac{\sigma_F^{-2}}{\sigma_{\phi\circ}^{-2}+\sigma_F^{-2}}\frac{(T+\tau-T')^2}{\tau^2}\right)
\end{aligned}$$

In the last step we introduce $\sigma_{\phi\circ} \equiv \frac{\sqrt{2}}{A\sqrt{N}}$, which is statistical error of the phase in the center of gravity frame from the 5-parameter fit alone. Analyzing eq.28, we find:

- No matter what parameters are, we always get improvement: $\sigma_\omega^{-2} \geq \sigma_{\omega\circ}^{-2}$.
- If $T' = T+\tau$, i.e. if phase is known at the time one lifetime after beginning of histogram, there is no improvement: $\sigma_\omega^{-2} = \sigma_{\omega\circ}^{-2}$. This is reasonable, because in this case there is no frequency–phase correlation and hence ω and ϕ are completely statistically independent.
- If phase is known at $t = T'$ with precision much better than $\sigma_{\phi\circ}$ (i.e. $\sigma_F^{-2} \gg \sigma_{\phi\circ}^{-2}$) then eq.28 for σ_ω^{-2} reduces to eq.23 for a fixed phase case: $\sigma_\omega^{-2} = \sigma_{\omega\circ}^{-2}[1+(T/\tau+1-T'/\tau)^2]$. For complete match, recall that $T' = 0$ in eq.23 due to the choice of time origin.
- For the opposite case when $\sigma_F^{-2} \ll \sigma_{\phi\circ}^{-2}$, there is no significant improvement for a practical value of T' such as $|T+\tau-T'| \sim \tau$: $\quad \sigma_\omega^{-2} \approx \sigma_{\omega\circ}^{-2}$.

As it's obvious from discussion above, figure of merit for σ_F is $\sigma_{\phi\circ}$. The goal of the current muon (g–2) experiment at BNL is to measure the muon spin (g–2) precession frequency at sub 10^{-6} level of accuracy. Thus

$$\sigma_{\phi\circ} = \sigma_{\omega\circ} \cdot \tau = \frac{\sigma_{\omega\circ}}{\omega} \cdot (\omega\tau) \sim 10^{-6} \cdot 100 = 10^{-4} \tag{29}$$

i.e. 0.1 mrad. Note that anglular displacement of muon spin with respect to its momentum in the muon storage ring by 0.1 mrad occurs in $(10^{-4}/2\pi) \times 4.365\,\mu$s = 69 ps, which corresponds to 30 cm/ns $\times$ 69 ps = 2.1 cm of muon (or decay electron) motion. This is the level of precision that we have to achieve in order to improve σ_ω by independent knowledge of the phase. To that level of precision we have to know average angle of polarization of muons injected into the storage ring, beam injection time and/or position of beam center at given time, average length or travel time of decay electrons from point of decay to calarimeter, etc. Each of these seems very challenging if possible at all. Therefore it was regarded as unpractical to persuade such a project in the current muon (g–2) experiment.

6 Appendix I

Consider equation for ellipsoid of errors

$$\sum_{i=1}^{L}\sum_{j=1}^{L} a_{ij}\ x_i\ x_j = 1 \tag{30}$$

with $a_{ij} = a_{ji},\ i = 1, 2, ..., L,\ j = 1, 2, ..., L$. Denote $L \times L$ matrix of parameters a_{ij} as $\mathcal{A}$. Statistical error for some k^{th} parameter is defined as an absolute value of extreme x_k coordinate of the ellipsoid defined by eq.30: $\sigma_k = x_k(max) = |x_k(min)|$. At these points $dx_k \equiv 0$ and we can use this to find σ_k. Take differencial from eq.30:

$$\sum_{i=1}^{L}\sum_{j=1}^{L} a_{ij}\ dx_i\ x_j + \sum_{i=1}^{L}\sum_{j=1}^{L} a_{ij}\ x_i\ dx_j = 2\ \sum_{i=1}^{L} dx_i \sum_{j=1}^{L} a_{ij}\ x_j = 0\ ,$$

$$\text{or} \qquad dx_k \sum_{j=1}^{L} a_{kj}\ x_j + \sum_{i \neq k} dx_i \sum_{j=1}^{L} a_{ij}\ x_j = 0 \tag{31}$$

For simplicity we consider the case of $k = 1$ first. Then eq.31 reads:

$$dx_1 \sum_{j=1}^{L} a_{1j}\ x_j + \sum_{i=2}^{L} dx_i \sum_{j=1}^{L} a_{ij}\ x_j = 0 \tag{32}$$

To ensure $dx_1 = 0$, we must vanish coefficients at all dx_i, $i \neq 1$. That leads to a system of $(L-1)$ linear equations:

$$\sum_{j=1}^{L} a_{ij}\ x_j = 0 \quad (i \neq 1)\ ,\ \text{or} \qquad \sum_{j=2}^{L} a_{ij}\ x_j = -a_{i1}\ x_1 \quad (i \neq 1) \tag{33}$$

We introduce matrix $\mathcal{A}'$, which can be constructed from matrix $\mathcal{A}$ by erasing first row and first column. Eg.33 can be written in matrix form as

$$\mathcal{A}' \times \boldsymbol{X'} = \boldsymbol{B} \cdot x_1 \tag{34}$$

where $\boldsymbol{X'}$ and $\boldsymbol{B}$ are $(L-1)-$dimension vectors of variables and constants, respectively. Explicitly:

$$X'_1 = x_2,\ X'_2 = x_3,\ \ \ldots\ ,\ X'_{L-1} = x_L \tag{35}$$

$$B_1 = -a_{21},\ B_2 = -a_{31},\ \ \ldots\ ,\ B_{L-1} = -a_{L1} \tag{36}$$

Using Cramer's rule, from eq.34 we can find X'_p in terms of x_1:

$$X'_p = \frac{\sum_{q=1}^{L-1} A'_{pq}\ B_q}{det(\mathcal{A}')}\ x_1\ , \qquad p, q = 1, ..., L-1 \tag{37}$$

where A'_{pq} and $det(\mathcal{A}')$ are cofactors of elements of matrix $\mathcal{A}'$ and its determinant, respectively. Now substitute eq.37 into eq.30 and find $x_1 = x_1(max) \equiv \sigma_1$. For convenience rewrite eq.30:

$$1 = \sum_{i=1}^{L}\sum_{j=1}^{L} a_{ij}\ x_i\ x_j = \sum_{i=2}^{L} x_i \sum_{j=1}^{L} a_{ij}\ x_j + x_1 \sum_{j=1}^{L} a_{1j}\ x_j \tag{38}$$

According to eq.33, the first term in the right side of eq.38 vanishes and we get:

$$\begin{aligned}\frac{1}{x_1} &= \sum_{j=1}^{L} a_{1j}\ x_j = a_{11}\ x_1 + \sum_{j=2}^{L} a_{1j}\ x_j = a_{11}\ x_1 + \sum_{p=1}^{L-1} a_{1(p+1)}\ X'_p = \\ &= x_1 \left(a_{11} + \frac{1}{det(\mathcal{A}')} \sum_{p=1}^{L-1} a_{1(p+1)} \sum_{q=1}^{L-1} A'_{pq}\ B_q \right)\end{aligned} \tag{39}$$

$$\text{then} \qquad \frac{det(\mathcal{A}')}{x_1^2} = a_{11}\ det(\mathcal{A}') + \sum_{p=1}^{L-1} a_{1(p+1)} \sum_{q=1}^{L-1} A'_{pq}\ B_q =$$

$$= a_{11}\ det(\mathcal{A}') - \sum_{p=1}^{L-1} a_{1(p+1)} \sum_{q=1}^{L-1} A'_{pq}\ a_{(q+1)1} \tag{40}$$

One can verify that the right side of eq.40 is nothing else but determinant of matrix $\mathcal{A}$ being calculated down the first column and then along the first row. Thus $x_1^{-2} \times det(\mathcal{A}') = det(\mathcal{A})$ and hence $\sigma_1 = \sqrt{det(\mathcal{A}')/det(\mathcal{A})}$. By definition of matrix $\mathcal{A}'$ as been constructed from matrix $\mathcal{A}$ by erasing its first row and first column, $det(\mathcal{A}')$ is equal to the minor M_{11} of $\mathcal{A}$. Thus

$$\sigma_1 = \sqrt{\frac{M_{11}}{det(\mathcal{A})}} \tag{41}$$

More general case of arbitrary k in eq.31 can be brought to the case of $k = 1$ by redefinition of k^{th} parameter into 1^{st}. Technically it can be done by $(k-1)$ cyclic permutations $(k-1 \to k,\ k-2 \to k-1, ..., 1 \to 2,\ k \to 1)$ of raws and similar $(k-1)$ cyclic permutations of columns of error matrix $\mathcal{A}$, which transforms $\mathcal{A}$ into some $\mathcal{A}^\star$. Since total number of permutations, $(2k-2)$, is even, determinants of initial and transformed error matrixes are equal. Furthermore, minors M_{kk} of initial and $M_{11}^\star$ of transformed error matrixes are also equal as being determinants of identical matrixes due to cyclicity of permutations, which preserve the order of remaining raws and columns. Therefore $\sigma_k = \sigma_1^\star = \sqrt{M_{11}^\star/det(\mathcal{A}^\star)} = \sqrt{M_{kk}/det(\mathcal{A})}$. Thus

$$\sigma_k = \sqrt{\frac{M_{kk}}{det(\mathcal{A})}} \tag{42}$$

for arbitrary k.

7 Appendix II

Here we give calculation of integrals, which we use for evaluation of statistical errors for the muon g–2 5-parameter fit.

$$\int f_\circ(t)dt = \int_T^\infty N_\circ\, e^{-t/\tau}\,[\,1 + A\,\cos(\omega t + \phi)\,]\,dt \approx N_\circ \int_T^\infty e^{-t/\tau} dt = N_\circ \tau\, e^{-T/\tau}$$

$$a_{11} = \frac{N\,e^{T/\tau}}{N_\circ \tau} \int_T^\infty \frac{1}{N_\circ} e^{-t/\tau}\,[\,1 + A\,\cos(\omega t + \phi)\,]\,dt \approx \frac{N}{N_\circ^2}$$

$$a_{12} = a_{21} = \frac{N\,e^{T/\tau}}{N_\circ \tau} \int_T^\infty \frac{t}{\tau^2}\, e^{-t/\tau}\,[\,1 + A\,\cos(\omega t + \phi)\,]\,dt \approx \frac{N}{N_\circ \tau}\,(T/\tau + 1)$$

$$
\begin{aligned}
a_{22} &= \frac{N\,e^{T/\tau}}{N_\circ \tau} \int_T^\infty N_\circ \frac{t^2}{\tau^4}\, e^{-t/\tau} \left[\, 1 + A\, \cos(\omega t + \phi)\, \right] dt \approx \frac{N}{\tau^2} \left[(T/\tau + 1)^2 + 1 \right] \\
a_{13} = a_{31} &= \frac{N\,e^{T/\tau}}{N_\circ \tau} \int_T^\infty e^{-t/\tau}\, \cos(\omega t + \phi)\, dt \approx 0 \\
a_{14} = a_{41} &= -\frac{N\,e^{T/\tau}}{N_\circ \tau} \int_T^\infty A\, t\, e^{-t/\tau}\, \sin(\omega t + \phi)\, dt \approx 0 \\
a_{15} = a_{51} &= -\frac{N\,e^{T/\tau}}{N_\circ \tau} \int_T^\infty A\, e^{-t/\tau}\, \sin(\omega t + \phi)\, dt \approx 0 \\
a_{23} = a_{32} &= \frac{N\,e^{T/\tau}}{N_\circ \tau} \int_T^\infty \frac{t}{\tau^2} e^{-t/\tau}\, \cos(\omega t + \phi)\, dt \approx 0 \\
a_{24} = a_{42} &= -\frac{N\,e^{T/\tau}}{N_\circ \tau} \int_T^\infty A\, \frac{t^2}{\tau^2}\, e^{-t/\tau}\, \sin(\omega t + \phi)\, dt \approx 0 \qquad (43) \\
a_{25} = a_{52} &= -\frac{N\,e^{T/\tau}}{N_\circ \tau} \int_T^\infty A\, \frac{t}{\tau^2} e^{-t/\tau}\, \sin(\omega t + \phi)\, dt \approx 0 \\
a_{33} &= \frac{N\,e^{T/\tau}}{N_\circ \tau} \int_T^\infty N_\circ e^{-t/\tau} \frac{\cos^2(\omega t + \phi)}{1 + A\, \cos(\omega t + \phi)}\, dt \approx \frac{N\,e^{T/\tau}}{2\tau} \int_T^\infty e^{-t/\tau} dt = \frac{N}{2} \\
a_{34} = a_{43} &= -\frac{N\,e^{T/\tau}}{N_\circ \tau} \int_T^\infty N_\circ A\, t\, e^{-t/\tau}\, \frac{\cos(\omega t + \phi) \sin(\omega t + \phi)}{1 + A\, \cos(\omega t + \phi)}\, dt \approx 0 \\
a_{35} = a_{53} &= -\frac{N\,e^{T/\tau}}{N_\circ \tau} \int_T^\infty N_\circ A\, e^{-t/\tau}\, \frac{\cos(\omega t + \phi) \sin(\omega t + \phi)}{1 + A\, \cos(\omega t + \phi)}\, dt \approx 0 \\
a_{44} &= \frac{N\,e^{T/\tau}}{N_\circ \tau} \int_T^\infty N_\circ A^2 t^2 e^{-t/\tau}\, \frac{\sin^2(\omega t + \phi)}{1 + A\, \cos(\omega t + \phi)}\, dt \approx \frac{N\, A^2\, \tau^2}{2} \left[(T/\tau + 1)^2 + 1 \right] \\
a_{45} = a_{54} &= \frac{N\,e^{T/\tau}}{N_\circ \tau} \int_T^\infty N_\circ A^2 t\, e^{-t/\tau}\, \frac{\sin^2(\omega t + \phi)}{1 + A\, \cos(\omega t + \phi)}\, dt \approx \frac{N\, A^2\, \tau}{2}\, (T/\tau + 1) \\
a_{55} &= \frac{N\,e^{T/\tau}}{N_\circ \tau} \int_T^\infty N_\circ A^2 e^{-t/\tau}\, \frac{\sin^2(\omega t + \phi)}{1 + A\, \cos(\omega t + \phi)}\, dt \approx \frac{N}{2}\, A^2
\end{aligned}
$$

In a course of calculation of these integrals we neglect terms of relative order of $\epsilon = (\omega\tau)^{-1} = 0.01$

References

[1] R. M. Carey et al., Phys. Rev. Lett. **82**, 1632 (1999);
H. N. Brown et al., Phys. Rev. **D62**, 091101 (2000);
H. N. Brown et al., Phys. Rev. Lett. **86**, 2227 (2001)

[2] F. James and M. Roos, “MINUIT, Function Minimization and Error Analysis”, CERN D506 (Long Writeup)