Some Things About $\Delta\Gamma$

R. Cahn (LBNL)

March 2001

1 Introduction

David Kirkby has raised a number of interesting questions that have in common that they depend on $\Delta\Gamma$, a quantity that we most often ignore in discussing B_d phenomena.

When we include the off diagonal matrix elements $\Delta\Gamma_{12}$ and $\Delta\Gamma_{21} = \Delta\Gamma_{12}^*$ the Schrödinger Equation becomes

$$i\frac{d}{dt}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} M - i\Gamma/2 & M_{12} - i\Gamma_{12}/2 \\ M_{12}^* - i\Gamma_{12}^*/2 & M - i\Gamma/2 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} \tag{1}$$

where a is the coefficient of $|B^0\rangle$ and b is the coefficient of $|\overline{B}^0\rangle$. Now finding the complex eigenvalues $\mu_\pm = M_\pm - i\Gamma_\pm$ we have

$$\mu_\pm = M - i\Gamma/2 \pm \sqrt{|M_{12}|^2 - |\Gamma_{12}|^2/4 - i\Re\Gamma_{12}M_{12}^*} \tag{2}$$

which we can conveniently expand as

$$\mu_\pm = M - i\Gamma/2 \pm |M_{12}|[1 - i\tfrac{1}{2}\Re\frac{\Gamma_{12}}{M_{12}}] \tag{3}$$

Now the usual definitions of p and q

$$\begin{aligned} |B_L\rangle &= p|B^0\rangle + q|\overline{B}^0\rangle \\ |B_H\rangle &= p|B^0\rangle - q|\overline{B}^0\rangle \end{aligned} \tag{4}$$

give

$$(M_{12} - i\Gamma_{12}/2)q = (\mu_- - M + i\Gamma/2)p \tag{5}$$

or

$$\begin{aligned} \frac{q}{p} &= -\frac{|M_{12}| - i\frac{1}{2}|M_{12}|\Re\frac{\Gamma_{12}}{M_{12}}}{M_{12} - i\Gamma_{12}/2} \\ &= -\frac{|M_{12}}{M_{12}}\left[1 - \tfrac{1}{2}\Im\frac{\Gamma_{12}}{M_{12}}\right] \end{aligned} \tag{6}$$

Now it is well known that $\Delta\Gamma$, which we see is

$$\Delta\Gamma = |M_{12}|\Re\frac{\Gamma_{12}}{M_{12}}, \tag{7}$$

is small. What is less obvious is that $\Im\Gamma_{12}/M_{12}$ is much smaller still.

The reason is the following. Both M_{12} and Γ_{12} arise from the box diagram changing B^0 to $\overline{B}^0$ through the exchange of two Ws. For M_{12} we insert all the charge 2/3 quarks. For Γ_{12} we are interested in transitions through physical states only, i.e. using u and c quarks only.

It is well known that the result for M_{12} is dominated by the t quark and is proportional to m_t^2 and to $(V_{tb}V_{td}^*)^2$. Now the diagrams for Γ_{12} can't know anything about the t quark mass, so that factor must be replaced by something else. It can't be the u or c quark masses, which are too small to matter. It is, in fact the b quark mass that comes in. This isn't too surprising, as it represents a phase space factor, appropriate for a real "decay." Since the light quark masses (and c is light here) don't appear, the various combinations of u and c will give us the combination

$$\begin{aligned}
&V_{ub}V_{ud}^*V_{ud}^*V_{ub} + V_{ub}V_{ud}^*V_{cd}^*V_{cb} \\
&\quad +V_{cb}V_{cd}^*V_{ud}^*V_{ub} + V_{cb}V_{cd}^*V_{cd}^*V_{cb} \quad = (V_{ub}V_{ud}^* + V_{cd}^*V_{cb})^2 \\
&\qquad = (-V_{tb}V_{td}^*)^2
\end{aligned} \tag{8}$$

So, to the extent that we drop light quark masses entirely, Γ_{12} has just the same phase as M_{12}! The corrections to this are of order m_c^2/m_b^2. Thus $\Im\Gamma_{12}/M_{12} \sim \mathcal{O}(m_c^2/m_t^2)$

A state that begins as B^0 at $t = 0$ evolves as

$$\begin{aligned}
|B^0{}_{phys}(t)\rangle &= \frac{1}{2p}[e^{-i\mu_- t}(p|B^0\rangle + q|\overline{B}^0\rangle) + e^{-i\mu_+ t}(p|B^0\rangle - q|\overline{B}^0\rangle)] \\
&= \tfrac{1}{2}e^{-iMt-\Gamma t/2}[(e^{i\frac{1}{2}\Delta mt+\frac{1}{4}\Delta\Gamma t} + e^{-i\frac{1}{2}\Delta mt-\frac{1}{4}\Delta\Gamma t})|B^0\rangle \\
&\quad +\frac{q}{p}(e^{i\frac{1}{2}\Delta mt+\frac{1}{4}\Delta\Gamma t} - e^{-i\frac{1}{2}\Delta mt-\frac{1}{4}\Delta\Gamma t})|\overline{B}^0\rangle] \\
&\equiv e^{-iMt-\Gamma t/2}[g_+(t)|B^0\rangle + \frac{q}{p}g_-(t)]|\overline{B}^0\rangle]
\end{aligned} \tag{9}$$

and similarly

$$|\overline{B}^0{}_{phys}(t)\rangle = \frac{1}{2q}[e^{-i\mu_- t}(p|B^0\rangle + q|\overline{B}^0\rangle) - e^{-i\mu_+ t}(p|B^0\rangle - q|\overline{B}^0\rangle)]$$

$$
\begin{aligned}
&= \tfrac{1}{2}e^{-iMt-\Gamma t/2}[\frac{p}{q}(e^{i\frac{1}{2}\Delta mt+\frac{1}{4}\Delta\Gamma t} - e^{-i\frac{1}{2}\Delta mt-\frac{1}{4}\Delta\Gamma t})|B^0\rangle \\
&\quad +(e^{i\frac{1}{2}\Delta mt+\frac{1}{4}\Delta\Gamma t} + e^{-i\frac{1}{2}\Delta mt-\frac{1}{4}\Delta\Gamma t})|\overline{B}^0\rangle] \\
&\equiv e^{-iMt-\Gamma t/2}[\frac{p}{q}g_-(t)|B^0\rangle + g_+(t)]|\overline{B}^0\rangle] \qquad (10)
\end{aligned}
$$

We see that

$$
\begin{aligned}
|g_+(t)|^2 &= \tfrac{1}{2}[\cosh\Delta\Gamma t/2 + \cos\Delta mt] \\
|g_-(t)|^2 &= \tfrac{1}{2}[\cosh\Delta\Gamma t/2 - \cos\Delta mt] \qquad (11)
\end{aligned}
$$

2 Decay Rates to Flavor Eigenstates

If we make the approximation that all final states have definite flavor, that is that each can be reached only from B^0 or from $\overline{B}^0$ we compute

$$
\begin{aligned}
|\langle B^0|B^0(t)\rangle|^2 &= \tfrac{1}{2}e^{-\Gamma|t|}[\cosh(\Delta\Gamma t/2) + \cos(\Delta mt)] \\
|\langle \overline{B}^0|B^0(t)\rangle|^2 &= \tfrac{1}{2}e^{-\Gamma|t|}|\frac{q}{p}|^2[\cosh(\Delta\Gamma t/2) - \cos(\Delta mt)] \\
|\langle \overline{B}^0|\overline{B}^0(t)\rangle|^2 &= \tfrac{1}{2}e^{-\Gamma|t|}[\cosh(\Delta\Gamma t/2) + \cos(\Delta mt)] \\
|\langle B^0|\overline{B}^0(t)\rangle|^2 &= \tfrac{1}{2}e^{-\Gamma|t|}|\frac{p}{q}|^2[\cosh(\Delta\Gamma t/2) - \cos(\Delta mt)]
\end{aligned} \qquad (12)
$$

Since $\Delta\Gamma/\Gamma << 1$, it suffices to expand $\cosh(\Delta\Gamma t/2) \to 1 + \Gamma^2 t^2/8$. For the moment, let's take $|q/p| = 1$. Then we can write our distributions as

$$
\begin{aligned}
|\langle B^0|B^0(t)\rangle|^2 &= \frac{\Gamma}{8}\left[1+\frac{\Delta\Gamma^2}{4\Gamma^2}\right]^{-1} e^{-\Gamma|t|}[1+\frac{\Delta\Gamma^2 t^2}{8} + \cos(\Delta mt)] \\
|\langle \overline{B}^0|B^0(t)\rangle|^2 &= \frac{\Gamma}{8}\left[1+\frac{\Delta\Gamma^2}{4\Gamma^2}\right]^{-1} e^{-\Gamma|t|}|[1+\frac{\Delta\Gamma^2 t^2}{8} - \cos(\Delta mt)] \\
|\langle \overline{B}^0|\overline{B}^0(t)\rangle|^2 &= \frac{\Gamma}{8}\left[1+\frac{\Delta\Gamma^2}{4\Gamma^2}\right]^{-1} e^{-\Gamma|t|}[1+\frac{\Delta\Gamma^2 t^2}{8} + \cos(\Delta mt)]
\end{aligned}
$$

$$|\langle B^0|\overline{B}^0(t)\rangle|^2 = \frac{\Gamma}{8}\left[1+\frac{\Delta\Gamma^2}{4\Gamma^2}\right]^{-1} e^{-\Gamma|t|}[1+\frac{\Delta\Gamma^2 t^2}{8} - \cos(\Delta m t)] \tag{13}$$

This is now properly normalized over the interval $-\infty < t < \infty$. It is apparent that the variable of interest is, in fact, $\zeta = \Delta\Gamma^2/\Gamma^2$. The question we address is how well can we measure ζ with N events taking into account the dilution $D = 1 - 2w$, but ignoring the z resolution?

So we consider

$$\begin{aligned}
|\langle B^0|B^0(t)\rangle|^2 &= \frac{\Gamma}{8}\left[1+\frac{\zeta}{4}\right]^{-1} e^{-\Gamma|t|}[1+\frac{\zeta\Gamma^2 t^2}{8} + D\cos(\Delta m t)] \\
|\langle \overline{B}^0|B^0(t)\rangle|^2 &= \frac{\Gamma}{8}\left[1+\frac{\zeta}{4}\right]^{-1} e^{-\Gamma|t|}|[1+\frac{\zeta\Gamma^2 t^2}{8} - D\cos(\Delta m t)] \\
|\langle \overline{B}^0|\overline{B}^0(t)\rangle|^2 &= \frac{\Gamma}{8}\left[1+\frac{\zeta}{4}\right]^{-1} e^{-\Gamma|t|}[1+\frac{\zeta\Gamma^2 t^2}{8} + D\cos(\Delta m t)] \\
|\langle B^0|\overline{B}^0(t)\rangle|^2 &= \frac{\Gamma}{8}\left[1+\frac{\zeta}{4}\right]^{-1} e^{-\Gamma|t|}[1+\frac{\zeta\Gamma^2 t^2}{8} - D\cos(\Delta m t)]
\end{aligned} \tag{14}$$

It is straightforward to calculate the distributions at $\zeta = 0$

$$\frac{\Gamma}{8}e^{-\Gamma|t|}(1 \pm D\cos\Delta m t) \tag{15}$$

and the derivatives of the distributions with respect to ζ at $\zeta = 0$:

$$-\tfrac{1}{4}\frac{\Gamma}{8}e^{-\Gamma|t|}[1 \pm D\cos\Delta m t - \tfrac{1}{2}\Gamma^2 t^2] \tag{16}$$

from which we determine

$$\sigma(\zeta)^{-2} = 2N\int_{-\infty}^{\infty} dt\frac{\Gamma}{8}e^{-\Gamma|t|}\left\{\frac{[1+D\cos\Delta m t - \frac{1}{2}\Gamma^2 t^2]^2}{16(1+D\cos\Delta m t)} + \frac{[1-D\cos\Delta m t - \frac{1}{2}\Gamma^2 t^2]^2}{16(1-D\cos\Delta m t)}\right\} \tag{17}$$

Dropping the D term, thus giving a conservative estimate, we have

$$\sigma(\zeta)^{-2} = \frac{5N}{16} \tag{18}$$

What is important here is that it is the error on $\Delta\Gamma^2/\Gamma^2$ that goes as $1/\sqrt{N}$:

$$\sigma(\Delta\Gamma^2/\Gamma^2) = \frac{4}{\sqrt{5N}} \tag{19}$$

Thus if we are setting a limit on $\Delta\Gamma/\Gamma$, that limit will improve only as $1/N^{1/4}$.

3 Can a Limit Vary Only as $N^{-1/4}$?

This may appear to conflict with the well-established principle that errors diminish as $1/\sqrt{N}$. The paradox can be resolved simply enough. Consider fitting to a parameter a whose true value is $\overline{a}$. We are told that the asymptotic error can be computed from the PDF $f(a)$ as

$$\sigma(a)^{-2} = N \int dx \frac{1}{f} \left(\frac{\partial f}{\partial a}\right)^2 \tag{20}$$

Now consider the variable $b = a^\nu$. We compute the anticipated error in b from

$$\frac{\partial f}{\partial b} = \frac{1}{\nu \overline{a}^{\nu-1}} \frac{\partial f}{\partial a} \tag{21}$$

so

$$\sigma(b) = \nu \overline{a}^{\nu-1} \sigma(a) \tag{22}$$

Now if $\overline{a}$ approaches zero if the error in a is well-behaved, the error in b won't be. Thus we found above that the error was well-behaved for $\Delta\Gamma^2$. It therefore isn't for $\Delta\Gamma$. Suppose that $\overline{a}$ isn't truly zero, but is only very small. As long as the error is large compared to the true value $\overline{a}$ we might as well suppose $\overline{a} = 0$. However, with sufficient data (N large enough), the error will eventually become small compared to $\overline{a}$. At that point the errors in a and b will both have conventional $1/\sqrt{N}$ behavior.

4 Petrak's Proposal to Use Just $B \to J/\psi K_S$

In BaBar Note 472, Sibylle Petrak shows how lifetime measurements on the $B \to J/\psi K_S$ data can be used to measure $\Delta\Gamma$ itself. We strip away much of the discussion and seize on her point that this can actually be done without tagging.

We begin with the decay amplitude to $B \to J/\psi K_S$ a time t after the identification of a B^0 state:

$$\mathcal{A} = g_+(t)pA + g_-(t)q\overline{A} \tag{23}$$

where

$$\begin{aligned} g_+(t) &= e^{-i\mu_- t} + e^{-i\mu_+ t} \\ g_-(t) &= e^{-i\mu_- t} - e^{-i\mu_+ t} \end{aligned} \tag{24}$$

Thus

$$|\mathcal{A}|^2 = |pA|^2[|g_+|^2 + 2\Re g_+^* g_- \lambda + |\lambda|^2 |g_-|^2] \tag{25}$$

Whereas if we start with a $\overline{B}^0$ the roles of g_+ and g_- are reversed:

$$|\mathcal{A}|^2 = |pA|^2[|g_-|^2 + 2\Re g_-^* g_+ \lambda + |\lambda|^2 |g_+|^2] \tag{26}$$

Now consider the sum

$$|\mathcal{A}|^2 + |\mathcal{A}|^2 = |pA|^2 \left[(1 + |\lambda|^2)(|g_+|^2 + |g_-|^2) + 4\Re g_+^* g_- \Re\lambda\right] \tag{27}$$

In fact

$$\begin{aligned} |g_\pm|^2 &= \tfrac{1}{2}e^{-\Gamma t}[\cosh\tfrac{1}{2}\Delta\Gamma t \pm \cos\Delta m t] \\ g_+^* g_- &= \tfrac{1}{2}e^{-\Gamma t}[\sinh\tfrac{1}{2}\Delta\Gamma t + i\sin\Delta m t] \end{aligned} \tag{28}$$

Combining these results

$$|\mathcal{A}|^2 + |\mathcal{A}|^2 = |pA|^2 e^{-\Gamma t}\left[(1 + |\lambda|^2)\cosh\tfrac{1}{2}\Delta\Gamma t + 2\sinh\tfrac{1}{2}\Delta\Gamma t\, \Re\lambda\right] \tag{29}$$

Specializing only slightly, we set $\lambda = -e^{-2i\beta}$ and find

$$|\mathcal{A}|^2 + |\mathcal{A}|^2 = 2|pA|^2 e^{-\Gamma t}\left[\cosh\tfrac{1}{2}\Delta\Gamma t - \cos 2\beta \sinh\tfrac{1}{2}\Delta\Gamma t\right] \tag{30}$$

Now really we have $-\infty < t < \infty$ and with normalization good to order $\Delta\Gamma$ the PDF is

$$f = \tfrac{1}{2}\Gamma e^{-\Gamma|t|}[\cosh\tfrac{1}{2}\Delta\Gamma t - \cos 2\beta\ \sinh\tfrac{1}{2}\Delta\Gamma t] \tag{31}$$

In fact, again to order $\Delta\Gamma$ we have simply

$$f = \tfrac{1}{2}\Gamma e^{-\Gamma|t|}[1 - \cos 2\beta\, \tfrac{1}{2}\Delta\Gamma t] \tag{32}$$

From this we find the asymptotic error

$$\sigma(\Delta\Gamma/\Gamma) = \sqrt{\frac{2}{N\cos^2 2\beta}} \tag{33}$$

This formula gives results in perfect agreement with those in BaBar Note 472, evidently obtained somewhat more laboriously. Of course we really need to consider the correlation with $\cos 2\beta$

5 Correlated Errors

In the analysis that uses the decays to all the flavor eigenstates we can compute the covariance matrix as long as we ignore the Δz resolution. For simplicity we'll also ignore the separation into mixed and unmixed events, which would be useful for determining the resolution function. Since the distribution is even in t we can restrict our consideration to $t > 0$ and write for the PDF

$$f = \Gamma(1 - \tfrac{1}{4}\zeta + \tfrac{1}{8}\zeta\Gamma^2 t^2)e^{-\Gamma t} \tag{34}$$

From this we find for small ζ

$$\begin{aligned}
-\frac{\partial^2\mathcal{L}}{\partial\Gamma^2} &= N\int_0^\infty dt\Gamma e^{-\Gamma t}(1-\Gamma t)^2/\Gamma^2 = N/\Gamma^2 \\
-\frac{\partial^2\mathcal{L}}{\partial\Gamma\partial\zeta} &= N\int_0^\infty dt\Gamma e^{-\Gamma t}(1-\Gamma t)(-\tfrac{1}{4}+\tfrac{1}{8}\Gamma^2t^2)/\Gamma = -\tfrac{1}{2}N/\Gamma \\
-\frac{\partial^2\mathcal{L}}{\partial\zeta^2} &= N\int_0^\infty dt\Gamma e^{-\Gamma t}(-\tfrac{1}{4}+\tfrac{1}{8}\Gamma^2t^2)^2 = \frac{5}{16}N
\end{aligned} \tag{35}$$

The covariance matrix is

$$\frac{1}{N}\begin{pmatrix} 5\Gamma^2 & 8\Gamma \\ 8\Gamma & 16 \end{pmatrix} \tag{36}$$

so that the errors are

$$
\begin{aligned}
\frac{\sigma(\Gamma)}{\Gamma} &= \sqrt{\frac{5}{N}} \\
\sigma([\frac{\Delta\Gamma}{\Gamma}]^2) &= \frac{4}{\sqrt{N}}
\end{aligned}
\tag{37}
$$

and the correlation between Γ and ζ is

$$
\rho(\Gamma,\zeta) = \frac{8}{\sqrt{5}\sqrt{16}} = \frac{2}{\sqrt{5}} = 0.89 \tag{38}
$$

It is this large correlation that inflates the errors so much beyond what they would be in the absence of correlations.

When we find $\Delta\Gamma$ from the $B^0 \to J/\psi K_S$ events there is no correlation between Γ and $\Delta\Gamma$ because the PDF is even in t for $\Delta\Gamma = 0$ while $\Delta\Gamma$ enters as an odd function of t. There is an important correlation with $\sin 2\beta$.

From

$$
f = \tfrac{1}{2}\Gamma e^{-\Gamma|t|}[1 - \cos 2\beta\, \tfrac{1}{2}\Delta\Gamma t] \tag{39}
$$

we find

$$
\begin{aligned}
-\frac{\partial^2\mathcal{L}}{\partial\Delta\Gamma^2} &= \int_{-\infty}^{\infty} dt \tfrac{1}{2} e^{-\Gamma|t|}(\tfrac{1}{2}t\cos 2\beta)^2 = \tfrac{1}{2}N\cos^2 2\beta/\Gamma^2 \\
-\frac{\partial^2\mathcal{L}}{\partial\Delta\Gamma\cos 2\beta} &= \int_{-\infty}^{\infty} dt \tfrac{1}{2} e^{-\Gamma|t|}(\tfrac{1}{2}t\cos 2\beta)(\tfrac{1}{2}\Delta\Gamma t) = \tfrac{1}{2}N\cos 2\beta\frac{\Delta\Gamma}{\Gamma^2}
\end{aligned}
\tag{40}
$$

The remaining element of the matrix is derived from the error on $\sin 2\beta$

$$
\begin{aligned}
-\frac{\partial^2\mathcal{L}}{\partial(\cos 2\beta)^2} &= [\sigma(\cos 2\beta)]^{-2} \\
\sigma(\cos 2\beta) &= \tan 2\beta\ \sigma(\sin 2\beta)
\end{aligned}
\tag{41}
$$

Inverting the covariance matrix $[\partial^2\mathcal{L}/\partial x\partial y]$ we find

$$
\begin{aligned}
\frac{\sigma(\Delta\Gamma)}{\Gamma} &= \sqrt{\frac{2}{N\cos^2 2\beta}} \times \left(1 - \tfrac{1}{2}\tan^2 2\beta\frac{\Delta\Gamma^2}{\Gamma^2}N\sigma^2(\sin 2\beta)\right)^{-1/2} \\
&= \sqrt{\frac{2}{N\cos^2 2\beta}} \times \left(1 - \frac{3\tan^2 2\beta\,\Delta\Gamma^2}{2Q\,\Gamma^2}\right)^{-1/2}
\end{aligned}
\tag{42}
$$

where we used the rule of thumb for the error in $\sin 2\beta$.

A rough comparison between the two approaches can be made. Suppose the total sample of nonleptonic and semileptonic decays is used, with say 25,000 events. This would give

$$\sigma([\frac{\Delta\Gamma}{\Gamma}]^2) = 0.025 \tag{43}$$

Thus the one-σ limit on $\Delta\Gamma/\Gamma$ is 0.16.

For the $B^0 \to J/\psi K_S$ sample suppose we have 300 events. Then ignoring the error from that of $\sin 2\beta$ (or noting that $\tan^2 2\beta$ is small) we would find

$$\sigma(\frac{\Delta\Gamma}{\Gamma}) = \sqrt{\frac{2}{300}} = 0.08 \tag{44}$$

6 Changing Variables in the Likelihood Function

While Γ and $\zeta = (\Delta\Gamma/\Gamma)^2$ are the natural variables for this study, it is interesting to ask how to express the results in terms of the variables $\Delta\Gamma$ and Γ. This is interesting enough to justify a slight excursion.

Consider expanding the log-likelihood function around its maximum:

$$d\mathcal{L} = \frac{\partial^2 \mathcal{L}}{\partial x_i \partial x_j} dx_i dx_j \tag{45}$$

using the index summation convention.

Now consider a second parameterization in terms of $\overline{x}_i$

$$\overline{\mathcal{L}}(\overline{x}) = \mathcal{L}(x) \tag{46}$$

where we can regard the xs as functions of the $\overline{x}$s so

$$dx_k = \frac{\partial x_k}{\partial \overline{x}_j} d\overline{x}_j \tag{47}$$

using the index summation convention. In this way we find

$$\frac{\partial^2 \overline{\mathcal{L}}}{\partial \overline{x}_i \partial \overline{x}_j} = \frac{\partial^2 \mathcal{L}}{\partial x_k \partial x_l} \frac{\partial x_k}{\partial \overline{x}_i} \frac{\partial x_l}{\partial \overline{x}_j} \tag{48}$$

which we rewrite as

$$\overline{L}_{lm} = L_{ij} U_{il} U_{jm} = U^T_{li} L_{ij} U_{jm} \tag{49}$$

Now the object of real interest is the covariance matrix, the inverse of L or $\overline{L}$.

$$\overline{L}^{-1} = U^{-1} L^{-1} U^{T^{-1}} \tag{50}$$

The meaning of elements of L^{-1} is

$$\begin{aligned} L^{-1}_{ii} &= \sigma^2(x_i) \\ L^{-1}_{ij} &= \sigma(x_i)\sigma(x_j)\rho_{ij} \end{aligned} \tag{51}$$

We can use this formalism to demonstrate the obvious feature that the uncertainty in, say, x_1, doesn't depend on the choice of the other parameters. Suppose $x_1(\overline{x}) = \overline{x}_1$. Then we can write

$$T = \begin{pmatrix} 1 & 0 \\ 0 & \tilde{T} \end{pmatrix} \tag{52}$$

where $\tilde{T}$ is a $(n-1)\times(n-1)$ matrix and the zeros are rows and columns of zeros. Similarly, write

$$L^{-1} = \begin{pmatrix} L^{-1}_{11} & L^{-1}_{1a} \\ L^{-1}_{a1} & L^{-1}_{ab} \end{pmatrix} \tag{53}$$

Then

$$\begin{aligned} \overline{L}^{-1} &= \begin{pmatrix} 1 & 0 \\ 0 & \tilde{T}^{-1} \end{pmatrix} \begin{pmatrix} L^{-1}_{11} & L^{-1}_{1a} \\ L^{-1}_{a1} & L^{-1}_{ab} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & (\tilde{T}^T)^{-1} \end{pmatrix} \\ &= \begin{pmatrix} L^{-1}_{11} & L^{-1}_{1a}(\tilde{T}^T)^{-1}_{ab} \\ \tilde{T}_{ca} L^{-1}_{a1} & \tilde{T}_{ca} L^{-1}_{ad} (\tilde{T}^T)^{-1}_{db} \end{pmatrix} \end{aligned} \tag{54}$$

In particular, we see that $\overline{L}^{-1}_{11} = L^{-1}_{11}$, showing that σ_1 is unchanged in going from $x_1, x_2, x_3 \ldots$ to $x_1, \overline{x}_2, \overline{x}_3 \ldots$.

Now suppose there are only two parameters, and still $x_1 = \overline{x}_1$. Now

$$T = \begin{pmatrix} 1 & 0 \\ \partial x_2/\partial \overline{x}_1 & \partial x_2/\partial \overline{x}_2 \end{pmatrix}; \qquad T^{-1} = \begin{pmatrix} 1 & 0 \\ -\frac{\partial x_2/\partial \overline{x}_1}{\partial x_2/\partial \overline{x}_2} & \frac{1}{\partial x_2/\partial \overline{x}_2} \end{pmatrix} \tag{55}$$

and

$$L^{-1} = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_{12} \\ \sigma_1\sigma_2\rho_{12} & \sigma_2^2 \end{pmatrix} \tag{56}$$

Combining these we can find the uncertainty in $\overline{x}_2$ and the correlation between $\overline{x}_1$ and $\overline{x}_2$:

$$\begin{aligned} \overline{\sigma}_2{}^2 &= \left(\frac{\partial x_2/\partial\overline{x}_1}{\partial x_2/\partial\overline{x}_2}\right)^2 \sigma_1^2 - 2\frac{\partial x_2/\partial\overline{x}_1}{(\partial x_2/\partial\overline{x}_2)^2}\sigma_1\sigma_2\rho_{12}\left(\frac{1}{\partial x_2/\partial\overline{x}_2}\right)^2\sigma_2^2 \\ \overline{\rho}_{12} &= \frac{\sigma_1/\overline{\sigma}_2}{\partial x_2/\partial\overline{x}_2}\rho_{12} - \frac{\sigma_1}{\overline{\sigma}_2}\frac{\partial x_2/\partial\overline{x}_1}{\partial x_2/\partial\overline{x}_2} \end{aligned} \tag{57}$$

Applying this to the case at hand,

$$\begin{aligned} x_1 = \Gamma, &\qquad x_2 = \zeta = (\Delta\Gamma/\Gamma)^2 \\ \overline{x}_1 = \Gamma, &\qquad \overline{x}_2 = \zeta = (\Delta\Gamma/\Gamma)^2 \\ \sigma_1 = \sqrt{\tfrac{5}{N}}\Gamma, &\qquad \sigma_2 = \frac{4}{\sqrt{N}} \\ \rho_{12} = \tfrac{2}{\sqrt{5}}, & \\ (\partial x_2/\partial\overline{x}_1) = -2\Delta\Gamma^2/\Gamma^3, &\quad (\partial x_2/\partial\overline{x}_2) = 2\Delta\Gamma/\Gamma^2 \end{aligned} \tag{58}$$

Combining these we find

$$\begin{aligned} \overline{\sigma}_2 &= \frac{2\Gamma^2}{\Delta\Gamma}\frac{1}{\sqrt{N}}\left[1 + 2\frac{\Delta\Gamma^2}{\Gamma^2} + \frac{5}{4}\frac{\Delta\Gamma^4}{\Gamma^4}\right]^{1/2} \\ \overline{\rho}_{12} &= \frac{4}{\sqrt{5}}[1 + \frac{5}{4}\frac{\Delta\Gamma^2}{\Gamma^2}]\left[1 + 2\frac{\Delta\Gamma^2}{\Gamma^2} + \frac{5}{4}\frac{\Delta\Gamma^4}{\Gamma^4}\right]^{-1/2} \\ &\approx \frac{4}{\sqrt{5}}[1 + \frac{1}{4}\frac{\Delta\Gamma^2}{\Gamma^2}] \end{aligned} \tag{59}$$

Actually we did not consistently kept higher order terms in $\Delta\Gamma/\Gamma$ in the initial derivation of $\sigma(\Gamma), \sigma(\zeta)$, and $\rho_{\Gamma,\zeta}$, so the higher order terms here are only indicative. We also set $D = 0$ in that earlier derivation.

7 CPT Violation

Belle reported [hep-ex/0011090 v3] limits on CPT violation using the formalism of T. D. Lee and C. S. Wu [*Ann. Rev. Nucl. Part. Phys.* **16**, 551 (1966)] and/or Lee and Wolfenstein [*Phys. Rev.* **138**, B1490 (1965)].

The correspondence between the usual formalism and this one is

$$\begin{pmatrix} \mu_{11} & \mu_{12} \\ \mu_{21} & \mu_{22} \end{pmatrix} = \begin{pmatrix} -iD + E_3 & E_1 - iE_2 \\ E + iE_2 & -iD - E_3 \end{pmatrix} \tag{60}$$

where

$$\begin{aligned} \mu_{11} &= M_{11} - i\tfrac{1}{2}\Gamma_{11} \\ \mu_{12} &= M_{12} - i\tfrac{1}{2}\Gamma_{12} \\ \mu_{21} &= M_{12}^* - i\tfrac{1}{2}\Gamma_{12}^* \\ \mu_{22} &= M_{22} - i\tfrac{1}{2}\Gamma_{22} \end{aligned} \tag{61}$$

The relation between μ_{12} and μ_{21} follows from the hermiticity of the Hamiltonian, i.e. from unitarity. Either CP or CPT invariance would make $\mu_{11} = \mu_{22}$, but we do not make this assumption here.

Some pseudo-angular variables are defined through

$$\begin{aligned} E^2 &= E_1^2 + E_2^2 + E_3^2 \\ E_3 &= E\cos\theta \\ E_1 &= E\sin\theta\,\cos\theta \\ E_2 &= E\sin\theta\,\sin\theta \end{aligned} \tag{62}$$

These angular variables are not generally real! Still, the usual trigonometric identities are valid, since they are simply algebraic relations among exponential functions that are true for real and complex arguments.

Explicitly

$$\begin{aligned} D &= i\tfrac{1}{2}\mu_{11} + \mu_{22} \\ E_1 &= \tfrac{1}{2}(\mu_{12} + \mu_{21}) \\ E_2 &= i\tfrac{1}{2}(\mu_{12} - \mu_{21}) \\ E_3 &= \tfrac{1}{2}(\mu_{11} - \mu_{22}) \\ E^2 &= \mu_{12}\mu_{21} + \tfrac{1}{4}((\mu_{11} - \mu_{22})^2 \end{aligned} \tag{63}$$

Now let us find the eigenenergies, $\mu_\pm$. Directly, we see

$$\mu_\pm = -iD \pm \sqrt{E_1^2 + E_2^2 + E_3^2} \tag{64}$$

We can make (unnormalized) eigenstates

$$\begin{aligned} |B_H\rangle &= |B^0\rangle + r_+|\overline{B}^0\rangle \\ |B_L\rangle &= |B^0\rangle + r_-|\overline{B}^0\rangle \end{aligned} \tag{65}$$

with

$$\begin{aligned} r_\pm &= \frac{\pm E - E_3}{E_1 - iE_2} = \frac{\pm 1 - \cos\theta}{e^{-i\phi}\sin\theta} \\ r_+ &= \tan\tfrac{1}{2}\theta e^{i\phi} \\ r_- &= -\cot\tfrac{1}{2}\theta e^{i\phi} \end{aligned} \tag{66}$$

Now we find the time evolution of states beginning as B^0 and $\overline{B}^0$ starting with

$$\begin{aligned} |B^0\rangle &= \frac{r_-|B_H\rangle - r_+|B_L\rangle}{r_- - r_+} \\ |\overline{B}^0\rangle &= \frac{|B_H\rangle - |B_L\rangle}{r_+ - r_-} \end{aligned} \tag{67}$$

Inserting the time dependence of the mass eigenstates,

$$\begin{aligned} |B^0(t)\rangle &= \frac{r_-e^{-i\mu_+t} - r_+e^{-i\mu_-t}}{r_- - r_+}|B^0\rangle + \frac{r_-r_+}{r_- - r_+}(e^{-i\mu_+t} - e^{-i\mu_-t})|\overline{B}^0\rangle \\ |\overline{B}^0(t)\rangle &= \frac{(e^{-i\mu_+t} - e^{-i\mu_-t})}{r_+ - r_-}|B^0\rangle + \frac{r_+e^{-i\mu_+t} - r_-e^{-i\mu_-t}}{r_+ - r_-}|\overline{B}^0\rangle \end{aligned} \tag{68}$$

Doing some algebra, we find

$$\begin{aligned} |B^0(t)\rangle &= [h_+ + \cos\theta h_-]|B^0\rangle + \sin\theta e^{i\phi}h_-|\overline{B}^0\rangle \\ |\overline{B}^0(t)\rangle &= \sin\theta e^{-i\phi}h_-|B^0\rangle + [h_+ - \cos\theta\, h_-]|\overline{B}^0\rangle \end{aligned} \tag{69}$$

where

$$h_\pm = \tfrac{1}{2}[e^{-i\mu_+t} \pm e^{-i\mu_-t}] \tag{70}$$

Now

$$
\begin{aligned}
|\langle B^0|B^0(t)\rangle|^2 &= |h_+|^2 + 2\Re\cos\theta h_- h_+^* + |\cos^2\theta|^2|h_-|^2 \\
|\langle \overline{B}^0|\overline{B}^0(t)\rangle|^2 &= |h_+|^2 - 2\Re\cos\theta h_- h_+^* + |\cos^2\theta|^2|h_-|^2 \\
|\langle \overline{B}^0|B^0(t)\rangle|^2 &= |\sin^2\theta e^{2i\phi}||h_-|^2 \\
|\langle B^0|\overline{B}^0(t)\rangle|^2 &= |\sin^2\theta e^{-2i\phi}||h_-|^2
\end{aligned} \tag{71}
$$

We actually would like to see this in terms of more physical quantities. First note that

$$
\begin{aligned}
|h_\pm|^2 &= \tfrac{1}{2}e^{-\Gamma t}(\cosh\tfrac{1}{2}\Delta\Gamma t \pm \cos\Delta m t) \\
h_+^* h_- &= -\tfrac{1}{2}(\sinh\tfrac{1}{2}\Delta\Gamma t + i\sin\Delta m t)
\end{aligned} \tag{72}
$$

Thus

$$
\Re\cos\theta h_+^* h_- = -\tfrac{1}{2}(\sinh\tfrac{1}{2}\Delta\Gamma t\Re\cos\theta - \sin\Delta m t\Im\cos\theta) \tag{73}
$$

and

$$
\begin{aligned}
|\langle B^0|B^0(t)\rangle|^2 &= (\cosh\tfrac{1}{2}\Delta\Gamma t + \cos\Delta m t) - (\Re\cos\theta\sinh\tfrac{1}{2}\Delta\Gamma t - \Im\cos\theta\sin\Delta m t) \\
&\quad + |\cos^2\theta|^2(\cosh\tfrac{1}{2}\Delta\Gamma t - \cos\Delta m t) \\
|\langle \overline{B}^0|\overline{B}^0(t)\rangle|^2 &= (\cosh\tfrac{1}{2}\Delta\Gamma t + \cos\Delta m t) + (\Re\cos\theta\sinh\tfrac{1}{2}\Delta\Gamma t - \Im\cos\theta\sin\Delta m t) \\
&\quad + |\cos^2\theta|^2(\cosh\tfrac{1}{2}\Delta\Gamma t - \cos\Delta m t) \\
|\langle \overline{B}^0|B^0(t)\rangle|^2 &= |\sin^2\theta||e^{2i\phi}|(\cosh\tfrac{1}{2}\Delta\Gamma t - \cos\Delta m t) \\
|\langle B^0|\overline{B}^0(t)\rangle|^2 &= |\sin^2\theta||e^{-2i\phi}|(\cosh\tfrac{1}{2}\Delta\Gamma t - \cos\Delta m t)
\end{aligned} \tag{74}
$$

Now it is easy to see that

$$
\begin{aligned}
e^{2i\phi} &= \frac{\mu_{21}}{\mu_{12}} \\
|e^{2i\phi}| &= \left|\frac{M_{12}^* - i\Gamma_{12}^*/2}{M_{12} - i\Gamma_{12}/2}\right| \\
&\approx 1 - \Im\frac{\Gamma_{12}}{M_{12}}
\end{aligned} \tag{75}
$$

while

$$\begin{aligned}\cos\theta &= \frac{E_3}{E} = \frac{\frac{1}{2}(\mu_{11}-\mu_{22})}{\sqrt{\mu_{12}\mu_{21}+\frac{1}{4}(\mu_{11}-\mu_{22})^2}} \\ &\approx \frac{M_{11}-M_{22}-\frac{1}{2}i(\Gamma_{11}-\Gamma_{22})}{\Delta m_d} \end{aligned} \tag{76}$$

A reasonable approach would be to keep first order terms in Γ_{12}/M_{12} and in $\cos\theta$. This would give

$$\begin{aligned}
|\langle B^0|B^0(t)\rangle|^2 &= (\cosh\tfrac{1}{2}\Delta\Gamma t+\cos\Delta m t)-(\Re\cos\theta\sinh\tfrac{1}{2}\Delta\Gamma t-\Im\cos\theta\sin\Delta m t)\\
|\langle \overline{B}^0|\overline{B}^0(t)\rangle|^2 &= (\cosh\tfrac{1}{2}\Delta\Gamma t+\cos\Delta m t)+(\Re\cos\theta\sinh\tfrac{1}{2}\Delta\Gamma t-\Im\cos\theta\sin\Delta m t)\\
|\langle \overline{B}^0|B^0(t)\rangle|^2 &= (1-\Im(\Gamma_{12}/M_{12}))(\cosh\tfrac{1}{2}\Delta\Gamma t-\cos\Delta m t)\\
|\langle B^0|\overline{B}^0(t)\rangle|^2 &= (1+\Im(\Gamma_{12}/M_{12}))(\cosh\tfrac{1}{2}\Delta\Gamma t-\cos\Delta m t)
\end{aligned} \tag{77}$$

8 Can We Let $\Delta\Gamma^2$ Go Negative?

Since we expect $\Delta\Gamma$ to be very small, we can anticipate that a fit for $\Delta\Gamma^2$, which is appropriate when we look at the time dependence of the Breco events, might lead to a negative and thus unphysical value. To simplify our considerations, in this section we ignore tagging so that the only PDF is

$$f=\Gamma(1-\frac{\Delta\Gamma^2}{4\Gamma^2})e^{-\Gamma t}\cosh\frac{\Delta\Gamma}{2}t \tag{78}$$

Of course, if the likelihood is computed as a function of $\Delta\Gamma$ there will inevitably be a maximum (or minimum of the negative log-likelihood) somewhere. This follows simply from the likelihood being an even function of $\Delta\Gamma$. However, if there is a minimum at exactly $\Delta\Gamma=0$, this minimum is spurious, due only to the identity

$$\frac{\partial\mathcal{L}}{\partial\Delta\Gamma}=2\Delta\Gamma\frac{\partial\mathcal{L}}{\partial\Delta\Gamma^2} \tag{79}$$

If the log-likelihood is plotted properly as a function of $\Delta\Gamma^2$ for $\Delta\Gamma^2>0$, the minimum may occur just at the endpoint $\Delta\Gamma^2=0$, without the slope of the likelihood function vanishing. This suggests that the likelihood function really wants to choose $\Delta\Gamma^2 < 0$. This might be managed by analytically continuing:

$$\cosh \tfrac{1}{2}\Delta\Gamma t \to \cos \tfrac{1}{2}\Delta\Gamma t \tag{80}$$

This seems dangerous because now the PDF is negative for some values of t. However, if $|\Delta\Gamma|$ is small enough we can get away with it. In particular, if the largest value of t is t_{max}, the log-likelihood is analytic in the full $\Delta\Gamma^2$ plane, except for cuts on the negative real axis beginning at

$$\Delta\Gamma^2 = -\frac{\pi^2}{t_{max}^2} \tag{81}$$

With N events, the expected value of t_{max} is $\ln N/\Gamma$. Since the fluctuations that might push $\Delta\Gamma^2$ negative should decrease as $1/\sqrt{N}$, we see that for large data sets, the problem posed by the possibility of the PDF going negative is not important.

Assuming still that $\Delta\Gamma/\Gamma << 1$, it is convenient to expand in

$$\xi = \frac{\Delta\Gamma}{4\Gamma^2} \tag{82}$$

We then have

$$f = \Gamma(1-\xi)e^{-\Gamma t}\cosh\sqrt{\xi}\Gamma t \tag{83}$$

For a dataset $\{t_i\}$, the log-likelihood is

$$\begin{aligned}\mathcal{L} &= \sum_i [\ln\Gamma + \ln(1-\xi) - \Gamma t_i + \ln\cosh\sqrt{\xi}\Gamma t_i] \\ &= N[\ln\Gamma - \xi - \tfrac{1}{2}\xi^2 - \Gamma\langle t\rangle + \tfrac{1}{2}\xi\Gamma^2\langle t^2\rangle - \tfrac{1}{12}\xi^2\Gamma^4\langle t^4\rangle \ldots]\end{aligned} \tag{84}$$

Here $\langle t\rangle$ means the mean value of t for this dataset. The maximum of $\mathcal{L}$ occurs when

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial\Gamma} &= 0 \quad = \frac{1}{\Gamma} - \langle t\rangle + \xi\Gamma\langle t^2\rangle + \ldots \\ \frac{\partial\mathcal{L}}{\partial\xi} &= 0 \quad = -1 - \xi + \tfrac{1}{2}\Gamma\langle t^2\rangle - \tfrac{1}{6}\xi\Gamma^4\langle t^4\rangle + \ldots\end{aligned} \tag{85}$$

The actual moments of the distribution are

$$\begin{aligned}
\overline{t} &= \Gamma^{-1}(1+2\overline{\xi}) \\
\overline{t^2} &= 2\Gamma^{-2}(1+5\overline{\xi}) \\
\overline{t^3} &= 6\Gamma^{-3}(1+9\overline{\xi}) \\
\overline{t^4} &= 24\Gamma^{-4}(1+14\overline{\xi})
\end{aligned} \tag{86}$$

where $\overline{\xi}$ is the true value of $\Delta\Gamma^2/(4\Gamma^2)$.

In terms that are already of order ξ, we can use the approximate relations

$$\overline{t^N} = \langle t^N \rangle = N!\Gamma^{-N} \tag{87}$$

In this way we find

$$\begin{aligned}
\xi &= \frac{\langle t^2 \rangle}{2\langle t \rangle^2} - 1 \\
\frac{1}{\Gamma} &= (1-2\xi)\langle t \rangle
\end{aligned} \tag{88}$$

Now ultimately ξ should be positive, but due to fluctuations, we may find that the best fit occurs for $\xi < 0$. The mean value of ξ is just $\overline{\xi}$. The variance of ξ is calculated from

$$\sigma_\xi^2 = \left(\frac{\partial\xi}{\partial\langle t\rangle}\right)^2 \sigma_t^2 + 2\left(\frac{\partial\xi}{\partial\langle t\rangle}\right)\left(\frac{\partial\xi}{\partial\langle t^2\rangle}\right)\sigma_{t,t^2}^2 + \left(\frac{\partial\xi}{\partial\langle t^2\rangle}\right)^2 \sigma_{t^2}^2 \tag{89}$$

where, using the $\overline{\xi} = 0$ approximation

$$\begin{aligned}
N\sigma_t^2 &= \overline{t^2} - \overline{t}^2 = \Gamma^{-2} \\
N\sigma_{t,t^2}^2 &= \overline{t^3} - \overline{t}\,\overline{t^2} = 4\Gamma^{-3} \\
N\sigma_{t^2}^2 &= \overline{t^4} - \overline{t^2}^2 = 20\Gamma^{-4}
\end{aligned} \tag{90}$$

Combining all this, we find that ξ has a mean of $\overline{\xi}$ and a variance

$$\sigma_\xi^2 = 1/N \tag{91}$$

Altogether, we expect the distribution of ξ to be

$$\frac{dP}{d\xi} = \frac{1}{\sqrt{2\pi}} e^{-N(\xi-\overline{\xi})^2/2} \tag{92}$$

As an example, suppose $\Delta\Gamma = 0.2\Gamma$ so $\overline{\xi} = 0.01$ and suppose there are $N = 5000$ events. Then the probability of finding a negative ξ is the probability of a downward fluctuation by $0.01\sqrt{5000} = 0.71$ sigmas, i.e. about 25%. (In Shahrams's presentation at the April Collaboration Meeting I find 315 entries in the first bin out of a total of 1610 entries, or 20%, even though Shahram says the fraction is 15%.)

Of course, we could find the uncertainty in ξ by evaluating the covariance matrix as we did in a previous section. The earlier result,

$$\sigma(\frac{\Delta\Gamma^2}{\Gamma^2}) = \frac{4}{\sqrt{N}} \tag{93}$$

is identical to the one obtained here.

Where does this leave us? The following procedure ought to work and give an optimal result:

- Treat $\Delta\Gamma^2$ as the Minuit variable.
- Fix the code so that when $\Delta\Gamma^2 < 0$, cosh becomes cos. Of course the cosh actually occurs inside a convolution. This will result in slightly different arguments for the erfc function in the event $\Delta\Gamma^2 < 0$.
- Run Minuit normally.
- Pray silently.
- If $\Delta\Gamma^2 > 0$ and in a statistically significant way, stare in amazement.
- If $\Delta\Gamma^2 > 0$, but not in a statistically significant way, set an upper limit for 90 or 95% confidence.
- If $\Delta\Gamma^2 < 0$, but not in a statistically significant way, set an upper limit for 90 or 95% confidence.
- If $\Delta\Gamma^2 < 0$, but in a statistically significant way, figure out what went wrong.
- To set the limits, it is probably best to ignore what Minuit has to say about errors and instead run toy Monte Carlos. The limit will be given on $\Delta\Gamma$.

It is a mistake to run Minuit with $\Delta\Gamma$ as the variable. If the result gives $\Delta\Gamma > 0$ no harm is done. However $\Delta\Gamma = 0$ is not improbable. If we simply accept this result, we are throwing away information, namely how negative $\Delta\Gamma^2$ really wanted to be. This will result in a weaker limit on $\Delta\Gamma$.