PRAMANA — journal of physics
© Indian Academy of Sciences
Vol. 55, Nos 1 & 2
July & August 2000
pp. 207–217

CP violation: the past, the present and the future

A I SANDA
Department of Physics, Nagoya University Nagoya, Japan

Abstract. We have just entered a period during which we expect considerable progress toward understanding CP violation. Here we review what we have learnt so far, and what is to be expected in the near future. To do this we cover the foundation of CP violation at a level which can be understood by physicists who are not working in this field.

Keywords. B meson; CP violation; K meson; weak interaction; flavour physics.

PACS Nos 13.20; 13.25.Es; 13.20.He

1. Introduction

CP violation was discoverd in 1964 [1]. It was a very good year, 1964, both in theory and in experimental high energy physics. For the following discoveries and breakthroughs were made that year:

1. The Higgs mechanism for the *spontaneous* realization of a symmetry was first developed.
2. The quark model and the first elements of current algebra were put forward.
3. The charm quark was first introduced to establish quark-lepton symmetry.
4. SU(6) symmetry was proposed.
5. The first storage ring for e^+e^- collisions was built in Frascati.
6. The Ω^- baryon was found at Brookhaven National Laboratory.
7. CP violation was discovered at the same lab.

1.1 *Expectations for the next decade*

The 1964 list is a phenomenal one covering a wide range of particle physics – it is almost impossible to beat! If we restrict ourselves to a field of CP violation, however, we believe that the next few years will be just as interesting and rewarding. Covering the period of 1999 to 2006, here is a list of experiments which have and will provide new information in this field:

1. CPLEAR collaboration finished a detail study of CP violation in K decays.
2. Finally the issue of ϵ'/ϵ was settled, as the KTeV collaboration [2] confirmed the non-vanishing value obtained by the NA31 collaboration [3].
3. A new accelerator, the Φ factory at DAΦNE, has been completed and KLOE collaboration is about to take data on coherently produced $K_L - K_S$ pairs.
4. Both B factories at KEK and SLAC have been completed and Belle and Babar collaborations have started data taking at respective laboratories.
5. At Fermilab, the main ring injector is in operation and CDF will start a serious attack on CP violation studies.
6. CLEO III will be taking data at ever increasing CESR luminosities.
7. HERAB is in its final development stage.
8. LHCB has been approved and continues its development.

As you see from this list, the excitement has already started. So, we will learn a lot more about this field in the next few years! Below, I will cover some basics of CP violation to appreciate the anticipated discoveries.

1.2 *CP violation – as a key to understanding our existence*

The fact that we exist, according to Sakharov [5], is due to CP violation. Our planet is made out of matter and a block of anti-matter can not exist. In fact we have not found any evidence of anti-matter in the whole universe. The amount of anti-particle found in cosmic rays is consistent with those produced in high-energy collisions of cosmic ray in earth's atmosphere. If this universe originated from the big-bang, which creates equal amount of matter and anti-matter, the forces responsible for the expansion and cooling of our universe must violate particle–antiparticle symmetry. CP violation must exist!

On earth, our daily life is controlled by physics of elementary particles with temperture from (0–1000)°C. In the beginning of the universe, about 10^{-20} sec after the big-bang, the temperture was about $10^{20\circ}$C. Particles were too energetic to form protons and neutrons. So, the universe was not made out of atoms, but their constituents quarks, leptons, photons, W bosons, Z bosons, and gluons. Before that, the universe was even hotter and made out of yet more fundamental particles.

At what stage did the asymmetry between matter and anti-matter get generated? If we know the answer to this question, then we know which physics we are probing when we study CP violation. What we know for sure is that it is not the Standard Model of elementary particle physics. This is because in the Standard Model with Kobayashi–Maskawa ansatz, CP violation is just a parameter. We can not compute this parameter nor do we understand its origin.

2. Introduction to CP violation

Let us start with a hamiltonian [4]

$$\mathcal{H} = cH + c^* H^\dagger. \tag{1}$$

The second term is required by hermiticity of the hamiltonian. If we examine the content of $H^\dagger$, we quickly discover the relationship between H and $H^\dagger$. Under particle to antiparticle transformation, $H \to H^\dagger$. That is, we have

$$\mathrm{CP}^\dagger H \mathrm{CP} = H^\dagger.$$

This means that if the coefficient c is real, $\mathcal{H}$ is invariant under CP . We have established that CP violation originates from existence of phases in the hamiltonian.

But not all the phases are physical observable. For example, electric charge is defined to be real because we can never observe its phase. This can be traced back to the fact that measurements are done by counting number of particles, which satisfy given conditions.

When we just count number of particles, most of the phase information is lost! The probability of detecting a given configuration is proportional to the absolute square of the amplitude, as we learned in quantum mechanics. Then how do we observe phases? We have encountered similar situation in optics. An interference pattern is created by phase difference between light waves which pass through two different slits. It can be detected by measuring the intensity of light on a screen.

It works the same way in an elementary particle system. To observe phases by just counting number of particles, we must rely on interference of two different amplitudes. To establish CP violation, for example in particle and anti-particle decays:

$$P \to X + Y \qquad \text{vs.} \qquad \bar{P} \to \bar{X} + \bar{Y}. \tag{2}$$

There must be:

(A) at least two amplitudes, with about the same magnitude, which lead to the same decay, (remember, in our optics example, if one of the slit is so narrow that not much light can pass through it, then the interference pattern can not be seen.);

(B) these amplitudes must have non-vanishing relative phase.

These conditions are very hard to realize in most decays. That's why CP violation is difficult to see. The reason why it is seen in the K system is because of an unusual and interesting phenomenon called particle–antiparticle mixing.

3. CP violation in the K system

Consider decays

$$K^0 \to \pi^+\pi^- \qquad \text{and} \qquad \bar{K}^0 \to \pi^+\pi^-. \tag{3}$$

The fact that both K^0 and $\bar{K}^0$ decay to a same channel implies that K^0 and $\bar{K}^0$ states are not orthogonal. They are not eigenstates of the hamiltonian. Also, the fact that both K^0 and $\bar{K}^0$ decay to a same channel means that there exist a transition of a type:

$$K^0 \leftrightarrow \pi^+\pi^- \leftrightarrow \bar{K}^0. \tag{4}$$

The time dependence of a single particle at rest is given by the following Schrödinger equation:

$$|K_{1,2}\rangle = a(t)|K^0\rangle + b(t)|\bar{K}^0\rangle,$$

$$i\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = \begin{pmatrix} M & M_{12} \\ M_{21} & M \end{pmatrix}\begin{pmatrix} a(t) \\ b(t) \end{pmatrix}. \tag{5}$$

By writing the diagonal elements to be equal, we have assumed CPT symmetry. Now write this matrix as:

$$\begin{pmatrix} M & M_{12} \\ M_{21} & M \end{pmatrix} = M\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & M_{12} \\ M_{21} & 0 \end{pmatrix}. \tag{6}$$

The expression on the right hand side of eq. (6) shows that eigenstate is controlled by the off diagonal terms–no matter how small M_{12} and M_{21} are relative to M. In fact for the K meson system $(M_{12}/M) \sim 10^{-12}$. If CP invariance holds, M_{12} is real, as argued above, and $M_{12} = M_{21}$. This requires complete mixing:

$$|K_{\pm}\rangle = \frac{1}{\sqrt{2}}(|K^0\rangle \pm |\bar{K}^0\rangle). \tag{7}$$

So, for example, in $K_{\pm} \to \pi^+\pi^-$ decay, amplitudes for two process, $A(K \to \pi^+\pi^-)$ and $A(\bar{A} \to \pi^+\pi^-)$, contribute and they will interfere.

Let us now consider CP violation. Where can the relative phase be present? There are two different ways in which the relative phase creaps in:

(i) in M_{12}
(ii) in the decay amplitude $A(K^0 \to \pi^+\pi^-)$ and $A(\bar{K}^0 \to \pi^+\pi^-)$.

Now, we could have just as well discussed $\pi^0\pi^0$ states instead of its charged counterpart. If M_{12} is the only place the phase is present, CP violation in $K_L \to \pi^+\pi^-$ is same as $K_L \to \pi^0\pi^0$, whereas if M_{12} is real and amplitudes $A(K^0 \to \pi^+\pi^-)$ and $A(\bar{K}^0 \to \pi^+\pi^-)$ possessed a relative phase, we could have a situation in which CP violating decay is only seen in $K_L \to \pi^+\pi^-$ decay and not in the neutral decay. The former is called a superweak scenario (indirect CP violation) and the latter is called direct CP violation.

The non-vanishing ϵ' means that there is difference between two CP violating decays: $K_L \to \pi^+\pi^-$ and $K_L \to \pi^0\pi^0$. This can not happen if the phase in M_{12} is the only source of CP violation. Thus the superweak scenario is out!

4. KM ansatz

We stated, in quite general grounds, that CP violation originates from the phase in the lagrangian. How does a phase, and therefore CP violation, appear in the standard model? Do we put them in by hand? No, surprisingly, it appears quite naturally – in other words, its absence is quite unnatural!

In the standard model with the KM ansatz, the gauge bosons couple as:

$$\mathcal{L} = \begin{pmatrix} \bar{u}^m \\ \bar{c}^m \\ \bar{t}^m \end{pmatrix}_L^T \mathbf{V}\gamma^\mu \begin{pmatrix} d^m \\ s^m \\ b^m \end{pmatrix}_L W_\mu + \mathrm{h.c.}, \tag{8}$$

Note that a W boson exhange changes flavor ($\mathbf{V}$ is not diagonal) and there is no restriction on $\mathbf{V}$ except for the fact that it must be unitary. Still, there are plenty of phases in $\mathbf{V}$. The problem is, as stated before, that experiments can only count number of particles. In quantities that can be measured, most of the phase information is lost. In particular, the phase of external states are not measurable. So, we can adjust them to make the constants which appear in the lagrangian real. Make the following transformation which does not change result of any experiment.

$$\begin{pmatrix}\bar{u}^m\\ \bar{c}^m\\ \bar{t}^m\end{pmatrix}_L \rightarrow \begin{pmatrix}\mathrm{e}^{i\phi_u} & 0 & 0\\ 0 & \mathrm{e}^{i\phi_c} & 0\\ 0 & 0 & \mathrm{e}^{i\phi_t}\end{pmatrix}\begin{pmatrix}\bar{u}^m\\ \bar{c}^m\\ \bar{t}^m\end{pmatrix}_L$$
$$\begin{pmatrix}\bar{d}^m\\ \bar{s}^m\\ \bar{b}^m\end{pmatrix}_L \rightarrow \begin{pmatrix}\mathrm{e}^{i\phi_d} & 0 & 0\\ 0 & \mathrm{e}^{i\phi_s} & 0\\ 0 & 0 & \mathrm{e}^{i\phi_b}\end{pmatrix}\begin{pmatrix}\bar{d}^m\\ \bar{s}^m\\ \bar{b}^m\end{pmatrix}_L . \tag{9}$$

What happens if we tune quark phases to make $\mathbf{V}$ real? Note that the number of parameters in $\mathbf{V}$ rises quadratically and the number of adjustable phases increases linearly with number of generations. For two generations, there are enough phases we can adjust to make $\mathbf{V}$ real. But, for 3 generations, there is not enough free phases to make all $\mathbf{V}_{ij}$ real. For a system of three generation of quarks, there are 6 phases and we can tune only 5 of them to zero following the above proceedure. One is left! This phase may appear in physical observable.

Note that we did not add additional CP violating interaction, nor invent an additional reason for making sure that a phase remains an observable. The existence of CP violation is a natural consequence of the fact that Yukawa couplings are in general complex – such a phase exists if we do not set it equal to zero by hand.

Now, all we have to do is to figure out how to measure it.

5. Key discoveries in the B system

Nature is kind to us! If she did not choose certain value for the elements of $\mathbf{V}$, we would not be here talking about CP violation in B decays. Here are two key discoveries which made this field so exciting.

5.1 *Longevity of B mesons*

The key discovery, which is the origin of many interesting phenomena in B physics, is that B mesons live for long time. This discovery was made possible by the development of vertex detectors to study charm mesons in hadronic colliders. In a hadronic environment, where there are large numbers of background tracks, neutral particles which travel some distance before they decay to charged particles leave a signature – a secondary vertex. A secondary vertex which is separated by few hundred microns from the interaction point can be identified.

The vertex detector turned out to be ideal for study of B mesons, which have about the same lifetime as D mesons. Experiments at SLAC, MAC [6] and MARK II [7] collabora-

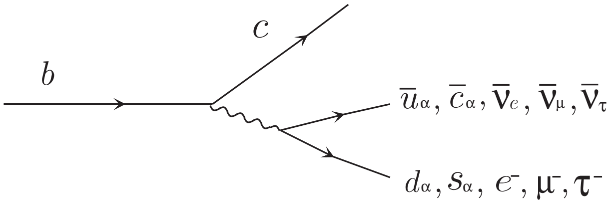

Figure 1. Feynman diagram which causes b-quark decays.

tions, have established the long lifetime of the B mesons. Their result was $(1.8 \pm 0.6 \pm 0.4)$ psec.

To appreciate the fact that this is a long time for a particle like the B meson, let us give a naive estimate of the lifetime. The total width of the b quark is given by just scaling the expression for the width of the μ meson:

$$\begin{aligned}\Gamma_\mu &= \frac{G_F^2}{192\pi^3} m_\mu^5 \\ &\Rightarrow \frac{G_F^2}{192\pi^3} m_b^5 |V_{bc}|^2 \times (2 \times 3 + 3). \qquad (10)\end{aligned}$$

There are altogether 9 channels (figure 1). For this rough estimate, the phase space difference due to finite quark masses is ignored. So, if $|V_{bc}| \sim 1$,

$$\tau \sim \tau_\mu \left(\frac{m_\mu}{m_b}\right)^5 \frac{1}{9} \sim 10^{-15} \mathrm{sec}. \qquad (11)$$

B's live 1000 times longer than our expectations! From this suppression, we deduce that

$$|\mathbf{V}_{cb}| \sim \frac{1}{30} \sim (\sin\theta_c)^2. \qquad (12)$$

5.2 $B^0 - \bar{B}^0$ *oscillations*

Now that we know that B mesons live for a long time, we might hope that it would do something interesting while it is alive. Indeed, $B - \bar{B}$ mixing was discovered by the ARGUS collaboration [8] by establishing the existence of same-sign dilepton events:

$$\begin{aligned}e^+e^- &\to \Upsilon(4S) \\ &\to B^0\bar{B}^0 \\ &\to \mu^\pm\mu^\pm + \text{anything}. \qquad (13)\end{aligned}$$

Of course if there is no mixing, we expect the leptons from $B - \bar{B}$ decay to have opposite signs. The rate for di-muon events can be characterized by

$$\frac{\Delta m}{\Gamma} = \frac{\text{lifetime}}{\text{mixing time}} \sim 0.7. \qquad (14)$$

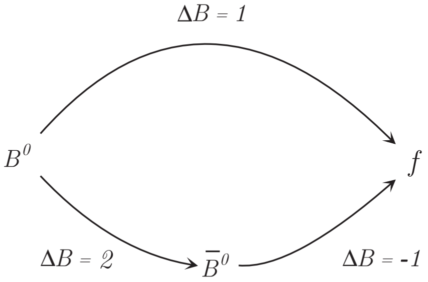

Figure 2. Two amplitudes which interfere to generate CP asymmetry in $B \to f$ decay.

When we were making predictions for this quantity, we had a notion that the top quark mass can not be more than 50 GeV. Since the rate for dimuon events goes like $(m_t/M_W)^4$, we failed to predict that the mixing could be seen. But, it should be noted that theorists have included the effect of mixing in their predictions [9].

6. CP asymmetry in B meson decays

The B meson system can be described in analogy with the K meson system. Here again as we discussed below, the interference of two amplitudes is crucial. Consider the problem of B decaying to a certain CP eigenstate f. Because of $B-\bar{B}$ mixing, there are two possible channels for $B \to f$ decay as shown in figure 2 [9].

A detailed calculation of this asymmetry gives a result

$$\frac{\Gamma(\bar{B} \to f) - \Gamma(B \to f)}{\Gamma(B \to f) + \Gamma(\bar{B} \to f)} = \mathrm{Im}\left(\frac{q}{p}\bar{\rho}\right)\sin(\Delta M t), \tag{15}$$

where

$$\frac{q}{p} = \left(\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}\right)^{\frac{1}{2}}, \tag{16}$$

and

$$\bar{\rho} = \frac{A(\bar{B} \to f)}{A(B \to f}. \tag{17}$$

Here the off diagonal mass matrix is written in two parts as shown in eq. (16). The diagrams which contribute to M_{12} and Γ_{12} are shown in figure 3.

These diagrams are proportional to the m_q^2 where m_q is the mass of the intermediate state quark. Note that only channels which B meson can actually decay can contribute to Γ_{12}. In particular, $t\bar{t}$ state can contribute only to M_{12} where there is no such constraint. This leads to the conclusion that

$$M_{12} \sim m_t^2 \quad \text{and} \quad \Gamma_{12} \sim m_c^2, \tag{18}$$

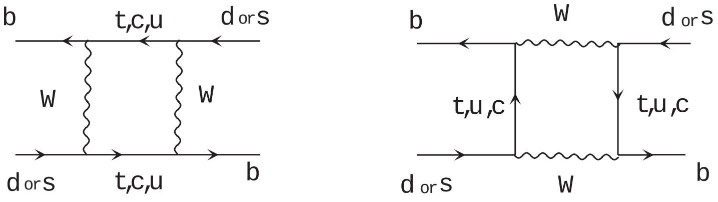

Figure 3. Feynman diagrams which are responsible for $B = 2$ transition.

and M_{12} dominates over Γ_{12}. If Γ_{12} can be neglected, ρ becomes very simple:

$$\begin{aligned}\frac{q}{p} &= \left(\frac{M_{12}^*}{M_{12}}\right)^{\frac{1}{2}} = \mathrm{e}^{-2i\arg M_{12}} \\ &= \mathrm{e}^{i2\arg(V_{td}V_{tb}^*)}. \end{aligned} \tag{19}$$

We shall now take a specific case in which $f = \psi K_S$ [10]. The diagram which contributes to ρ are shown in figure 4.

Each amplitude depends on several form factors describing how quarks form certain hadronic states. We do not know how to compute these form factor. But, for this special example, all form factors cancel when we take the ratio, and ρ is given by the KM matrix elements alone:

$$\begin{aligned}\bar{\rho} &= \frac{V_{cb}V_{cs}^*}{V_{cb}^*V_{cs}} \\ &= \mathrm{e}^{2i\arg(V_{cb}V_{cs}^*)}. \end{aligned} \tag{20}$$

So, the asymmetry defined in eq. (15) can be evaluated by

$$\mathrm{Im}\left(\frac{q}{p}\bar{\rho}\right) = \sin(2\phi_1). \tag{21}$$

where

$$\phi_1 = \arg(V_{cb}V_{cs}^*V_{td}V_{tb}^*). \tag{22}$$

What makes B system simple?

In our illustration above, we used $K \rightarrow \pi\pi \rightarrow \bar{K}$ as an example of a channel which causes $K - \bar{K}$ mixing. There are other intermediate states, and for the K system, M_{12} and Γ_{12} are comparable in size. This makes it difficult to predict q/p. But, for the B system $B \rightarrow t\bar{t} \rightarrow \bar{B}$, as shown in figure 3, by far dominates all the other intermediate states since $M_{12} \propto m_t^2$. This leads to a simple relation eq. (19). Because ψK_S is a pure isospin state, in addition to being a CP eigenstate, there is only an overall strong interaction phase which cancels when we take the ratio in ρ. So, the ψK_S asymmetry can be expressed purely in terms of the KM matrix elements.

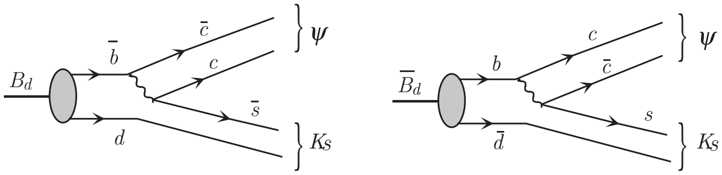

Figure 4. The diagram which contributes to ρ.

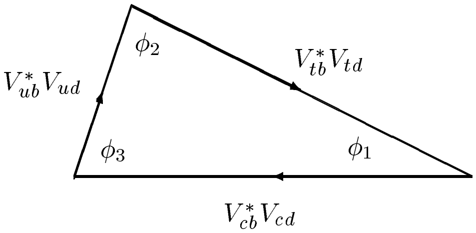

Figure 5. The unitarity triangle in the complex plane.

7. Unitarity triangle

One of the constraints from unitarity of KM matrix elements is:

$$V_{ub}^*V_{ud} + V_{cb}^*V_{cd} + V_{tb}^*V_{td} \;=\; 0\,. \tag{23}$$

This relation corresponds to a triangle in the complex plane, the so-called unitarity triangle and is shown in figure 5.

Its three inner angles is denoted by

$$\begin{aligned}\phi_1 &= \pi - \arg\left(\frac{V_{tb}^*V_{td}}{V_{cb}^*V_{cd}}\right),\\ \phi_2 &= \arg\left(\frac{V_{tb}^*V_{td}}{-V_{ub}^*V_{ud}}\right),\\ \phi_3 &= \arg\left(\frac{V_{ub}^*V_{ud}}{-V_{cb}^*V_{cd}}\right).\end{aligned} \tag{24}$$

This triangle takes a very simple form when we write the sides in terms of Wolfenstein representation of the KM matrix.

$$\mathbf{V} = \begin{pmatrix} 1-\frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta + \frac{i}{2}\eta\lambda^2) \\ -\lambda & 1-\frac{1}{2}\lambda^2 - i\eta A^2\lambda^4 & A\lambda^2(1+i\eta\lambda^2) \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix}. \tag{25}$$

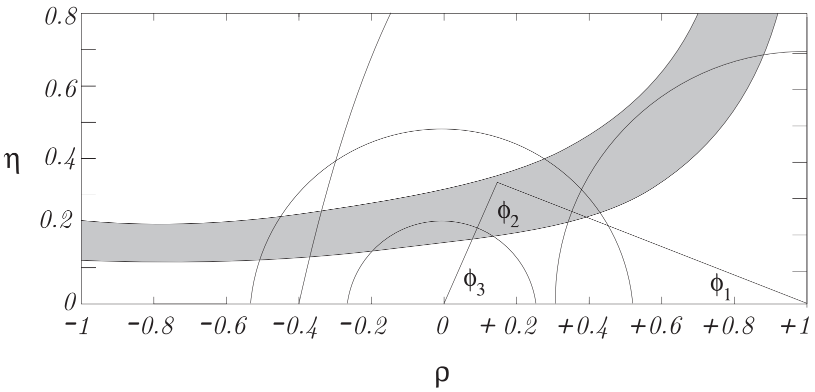

Figure 6. Experimentally allowed region for (ρ, η). The region within the two semi-circle follows from experimental measurement of the leptonic B decay. The region between two curves come from the constrain from ϵ. The width of the region reflects the theoretical ambiguity in the theoretical prediction. Reproduced with permission from Cambridge University Press.

Note that all three terms which appear in eq. (23) are proportional to $V_{cb}^* V_{cd} \sim A\lambda^3$. After this factor is scaled out, we have the triangle shown in figure 6.

We know something about this unitarity triangle. For example, study of leptonic B decays leads to information about the quark decay $b \to u\, e\, \nu$. The strength of this transition is given by $|V_{ub}| \sim \sqrt{\rho^2 + \eta^2}$ which in turn is related to the length of the vector defined by the point (ρ, η) on the $\rho - \eta$ plane. The experimental data on $|V_{ub}|$ defines a semi-circle shown in figure 6 [11].

The Nature chooses a value for (ρ, η) between the two semi-circle which reflect experimental measurement error. An explicit theoretical evaluation of ϵ, together with the experimental value for ϵ yield a curve on the $\rho - \eta$ plane. As mentioned above, the theoretical ambiguity in computing ϵ, due to strong interaction, allow us to give only a band on the $\rho - \eta$ plane. If the standard model of CP violation is correct, one of the edges of the unitarity triangle must lie in the overlap region shown in figure 6. This in turn implies that CP violation in $B \to \psi K_S$ is at least 20% and can be nearly as big as 100%.

8. Summary

CP violation in $B \to \psi K_S$ can be anywhere from 20% to 100%. This is very exciting. This is only one of the experimental observables related to the unitarity triangle shown in figure 6. In principle, there are six quantities associated with the triangle to be measured. Since three quantities define a triangle uniquely, there will be three consistency checks for the predictions of the standard model.

Even such a large asymmetry is very difficult to measure as B mesons lives for only about 10^{-12} sec. Not only that, all of above discussion assumes that we have a B or $\bar{B}$ beam. Actually, B and $\bar{B}$ are pair-produced from the decay of $\Upsilon(4S)$. So, we have to tag one B meson to determine the identity of the other one as illustrated in figure 7. Then the

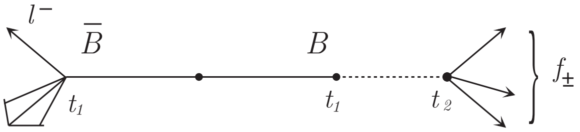

Figure 7. To get, for example, a B beam, we make sure that at time t_1 we have $\bar{B}$ by tagging it. The other unobserved partner must be a B at time t. So, setting t_1 to be the initial time, we observe the time evolution of the B state. Reproduced with permission from Cambridge University Press.

tag time is $t = 0$. We have to determine the time difference between the two decays. Note that the mass of $\Upsilon(4S)$ is only slightly above $2 \times m_B$.

This is both a curse and a blessing. Its a curse because in the rest frame of $\Upsilon(4S)$, B has so little kinetic energy that it moves only about 20 microns before it decays. It is impossible with today's technology to measure tracks with resolution much better than 20 micron. So, we can not do the experiment in the rest frame of $\Upsilon(4S)$. We have to build an asymmetric e^+e^- storage ring. At asymmetric B factories at KEK and SLAC, $\Upsilon(4S)$ is boosted so that B mesons travels average of about 200 microns before they decay. In the long run, it is a blessing because the $\Upsilon(4S)$ decay is very pure. It can be shown that $B - \bar{B}^*$ pair will result in different type of CP asymmetry. So, this mixture would have introduced another uncertainty.

Asymmetric B factories are now complete at KEK and at SLAC, and both laboratories are taking data. This will be a tight race to get at the asymmetry.

Acknowledgements

The author's research is supported by Grant-in Aid for Special Project Research (Physics of CP violation).

References

[1] A Pais, CP violation: the first 25 years, *CP Violation in particle and astrophysics* edited by J Tran Than Van (Edition Frontières, Gif-sur-Yvette, France, 1989)
[2] A Alavi-Harati *et al*, *Phys. Rev. Lett.* **83**, 22 (1999)
[3] NA31 collaboration: G D Barr *et al*, *Phys. Lett.* **B317**, 233 (1993)
[4] For a further introduction, and details, see I I Bigi and A I Sanda, *CP Violation* (Cambridge University Press, Cambridge, 1999)
[5] A D Sakharov, *JETP Lett.* **5**, 24 (1967)
[6] E Fernandez *et al*, *Phys. Rev. Lett.* **51**, 1022 (1983)
[7] N Lockyer *et al*, *Phys. Rev. Lett.* **51**, 1316 (1983)
[8] H Albrecht *et al*, *Phys. Lett.* **B192**, 245 (1987)
[9] A B Carter and A I Sanda, *Phys. Lett.* **23**, 1567 (1981)
[10] I I Bigi and A I Sanda, *Nucl. Phys.* **B193**, 85 (1981)
[11] Particle Data Group, *Europhys. J.* **C3**, 1 (1998)